MPHIL

Control of Soil Transmitted Helmiths

Boulton, Beth

Award date:
2020

Awarding institution:
University of Bath

Control of Soil Transmitted Helminths

submitted by

Bethan Boulton

for the degree of Master of Philosophy

of the

University of Bath

Department of Mathematical Sciences

November 2019

Summary

Helminths are macroparasites that cause many diseases, in people and animals, with a mode of transmission very different from that of most bacteria and viruses. The severity of the infection in a host is affected by the parasite burden of that specific host. The parasite burdens are extremely variable between individual hosts as a result of the transmission methods. The diseases caused by parasites are often difficult to treat effectively, and even more difficult to eradicate, largely because of the complex transmission methods. As such, models could be very helpful in determining best practices in the efficient treatment of these diseases.

The research presented here is an investigation into the spread of soil-transmitted helminths, through the use of metapopulation models, and the effects of simple control measures on the infections caused by them. Models often fall into two categories: fully stochastic models, which are computationally costly; or mean-field models, which cannot capture the detail of this variability. The methodology of this research is to formulate models that incorporate the variance in parasite burdens between hosts to study whether the variance would cause a significant difference in the recommended control, based on the optimisation and study of simple controls on the models.

The initial examination is intended to see how the inclusion of additional details specific to populations in multiple environments, such as the inclusion of the variance in parasite burdens between hosts, could be incorporated into the models and the effect this would have on the controls.

In this investigation, we found that simply including the variance in the models did not significantly improve the predictive power of the model in regards to control. Without additional knowledge of the underlying distribution of the parasite burdens, which the mean and variance serve to model, treatments could only be applied to the whole host population. This means that the effect of more individualised treatments, only on hosts that really need it, cannot be studied and the increased information of the distribution that the variance provides is not utilised. For these treatments on the entirety of the host population, we found that, comparing models which included parasite burden variability and those which did not, did not result in qualitatively different optimal controls. This was the case even when the control was optimised with the intent of minimising the variance as a priority. However, we also took steps to consider how treatments could be applied to only a portion of the overall host population being modelled. This was only possible when the mean and variance of the model were a good fit for the underlying distribution of parasite burdens, which was known as a result of empirical means. The hope is that this will provide a potential avenue for further study in this area.

On the other side of the investigation, we studied how the inclusion of parasite resistance to treatment may be modelled and the effects on the controls. In investigating this we found that the total parasite populations could posses properties that the standard model analysis would not show. The

most important of these was the potential that the basic model would be unreactive, which would imply that transient growth around an equilibrium was unlikely, yet when the sum of variables was considered transient growth would occur. These properties could then have significant effects on the control. This is due to the effectiveness of the control being measured against the change to the total population, rather than the individual variables.

It is hoped that this research may further the understanding of how variance may be incorporated into metapopulation models and the potential difficulties, and how parasite resistance to treatment can be included in models and its effects on treatment to improve prediction in treatments.

Contents

1 Introduction and Background **6**

1.1 Motivation and Question . . . 6
1.2 Macroparasite Background . . . 7
1.3 Previous Models of Macroparasites . . . 9
1.4 Metapopulation Modelling . . . 13
1.5 Summary . . . 15

2 Preliminary Models **16**

2.1 Moment Closure Approximation . . . 18
2.1.1 Model Examples . . . 19
2.2 Application of Treatment . . . 22
2.2.1 Treatment at Set Times . . . 22
2.2.2 Parasite Burden Dependent Treatment . . . 25
2.2.3 Burden Dependent Treatment at Fixed Times . . . 26
2.3 Attempts to Incorporate the Larval Stage . . . 29
2.3.1 Simulation Model . . . 29
2.3.2 Model Examples . . . 31
2.4 Discussion of Models . . . 38

3 Single Species of Parasite Models **39**

3.1 Model Basis . . . 39
3.1.1 Introducing Mean Field Models . . . 39
3.2 Individual Simulation Model . . . 41
3.3 Moment Closure on Variance . . . 44
3.3.1 Analysis of Model . . . 45
3.4 Direct Inclusion of Variance . . . 49
3.4.1 Analysis of the Model . . . 52
3.5 Conclusions . . . 57

4 Application of Control **59**

4.1 Basic Optimal Control . . . 59
4.1.1 Pontryagin's Principle . . . 60

4.1.2 Control Example Parameters . . . 61
4.1.3 Application to the Mean-Field Model . . . 62
4.1.4 Application to the Variance Inclusive Model . . . 66
4.1.5 Application to Isham Model . . . 70
4.2 Integer Controls . . . 74
4.2.1 Method 1 . . . 75
4.2.2 Method 2 . . . 76
4.3 Control as a Switching System . . . 80
4.3.1 Application to the Mean-Field Model . . . 81
4.3.2 Application to the Variance Inclusive Model . . . 82
4.3.3 Application to the Isham Model . . . 83
4.4 Discussion . . . 88

5 Resistant Parasite Models 89
5.1 Haploid Models . . . 89
5.1.1 Two Genotype Mean-Field Model . . . 91
5.1.2 Two Genotype Variance Inclusive Model . . . 92
5.2 Haploid Model Analysis . . . 92
5.2.1 Two Genotype Mean-Field Model Analysis . . . 93
5.2.2 Two Genotype Variance Inclusive Model Analysis . . . 95
5.3 Diploid Model . . . 101
5.3.1 Model Formulation . . . 102
5.3.2 Numerical Simulation . . . 106
5.4 Discussion . . . 111

6 Control on Resistant Models 112
6.1 Model Dynamics Under Treatment . . . 112
6.1.1 Haploid Models . . . 112
6.1.2 Diploid Model . . . 115
6.2 Continuous Optimal Control . . . 121
6.2.1 Haploid Models . . . 121
6.2.2 Diploid Model . . . 130
6.3 Discussion . . . 135

7 Integer Controls on Resistant Models 137
7.1 Haploid Mean-Field Model . . . 137
7.2 Haploid Variance Inclusive Model . . . 143
7.3 Diploid Model . . . 153
7.4 Discussion . . . 158

8 Conclusions and Discussion 160
8.1 Key Results . . . 160

8.2 Future Work . . . 163
8.2.1 Split Control on the Isham model . . . 163

Appendices **175**

A Chapter 2 Codes **176**
A.1 Isham Codes . . . 176
A.1.1 Basic Isham Model . . . 176
A.2 Larval Model Codes . . . 176

B Chapter 3 Codes **177**
B.1 Expected Number of Mated Females . . . 177
B.2 ODE Models . . . 177

C Chapter 4 Codes **178**
C.1 Bocop Files . . . 178
C.2 Sequentially Determined Integer Controls . . . 178
C.3 Brute Force Optimised Integer Controls . . . 179
C.4 Switching System Controls . . . 179

D Chapter 5 Codes **180**
D.1 Haploid Models . . . 180
D.2 Diploid Model . . . 180

E Chapter 6 Codes **181**
E.1 Haploid Model Codes . . . 181
E.2 Diploid Model Codes . . . 181

F Chapter 7 Codes **182**

G Additional Figures **183**

H Future Work Codes **185**

Chapter 1

Introduction and Background

1.1 Motivation and Question

This thesis presents a series of mathematical models to explore the impact of population variation on population control. To achieve this, the model examples are based on macroparasitic infections of hosts where parasite burden within individual hosts varies. To contextualise this system further, we consider the individual hosts to behave as a metapopulation system for the macroparasites which themselves can move between the hosts. The host-macroparasite system also lends itself to consider the impact of resistance on control strategies, since this is often observed in such systems.

A metapopulation is a group of populations from the same species that occupy spatially segregated patches but can interact through the movement of individuals between the patches [41, 31, 42, 40, 52]. In a metapopulation model, the dynamics of each sub-population are affected not only by the behaviour of individuals within that sub-population but also by the dynamics of the linked sub-populations through processes such as immigration. Metapopulations exist throughout nature, but they are often ignored for simplicity in favour of examining individual populations within a single environment [49, 80, 17]. In many cases, this may make sense. For example, if the link between sub-populations is very weak then it may be more appropriate to consider two distinct populations; by contrast, if the link is very strong then it might be more appropriate to join the two sub-populations together and treat them as a single population. When neither of these scenarios is appropriate, the use of a metapopulation structure allows us to explore the potential for distinct sub-populations to act as reservoirs to re-establish dwindling populations [44]. In other words, a metapopulation model allows for the environmental heterogeneity to affect the persistence and spread of a species. Based on this brief description, it is clear that significant impact may be observed on population persistence by using a metapopulation structure in preference to a single, homogeneous population model [45].

Macroparasites provide a good example of how such investigations may be useful. Although not traditionally considered metapopulations, macroparasites have complex life-cycles [25], discussed in greater detail in section 1.2, which allows each host to act as a separate environment for the para-

sites. As a result, the infection dynamics may be drastically different from those of microparasites, which many infection models have focused on.

Often mean-field models are used to examine the average parasite population within a host, which is not necessarily the most effective method. Parasite infections are rarely evenly distributed throughout host populations. As such, the predictive capacity of models that consider only the average parasite load of all individuals in a population is likely to be limited [37, 99]. Variation in infection status between hosts can have consequences for the transmission dynamics and the severity of the infection of individual hosts. Mean-field models, as the name suggests, are formulated to examine the average behaviour of a population. These mean-field models typically consider a single population, although they could also examine the mean of linked populations [7, 67]. In contrast, metapopulation models consider linked populations, either by modelling all individual sub-populations involved or by examining the probability of different environments being populated. These metapopulation models can be better at capturing the variability between sub-populations. However, for a large number of hosts, these can result in large systems and it can be difficult determining individual rate parameters for all hosts [41].

In this research, we formulate a methodology by which a greater representation of the distribution of parasites can be determined, without needing a fully stochastic model or a large deterministic system. We do this by considering models that are formulated to include the variance in the number of parasites between hosts. With these models, we will then study how the inclusion of the variance changes the way that control strategies are developed and use the models to predict the efficacy of these control strategies.

1.2 Macroparasite Background

Macroparasites are organisms that live within, or on, a host species at the host's expense and are large enough to be seen with the naked eye [72]. Macroparasites encompass different varieties such as helminths, which include worms and other internal parasites, and arthropods, such as ticks [9]. In this research, we consider helminth infections. To understand the modelling techniques needed to consider macroparasitic infections it is important to understand the crucial differences, primarily in the life cycle, between them and infections caused by microparasites, such as viruses and bacteria. In classic microparasitic infections [8, 53] the pathogens are able to multiply within the host, while macroparasites typically cannot [89]. Instead, macroparasites have complex life-cycles that involve early stage parasites, such as eggs, being passed out of a host to mature through various life stages before they are picked up by another host [50, 25]. A basic macroparasitic life-cycle is as follows:

1) Infective stage parasite is acquired by its host.

2) The parasite migrates to the region they infect within the host and mature.

3) Parasites reproduce within the host and produce eggs, which are passed out of the host.

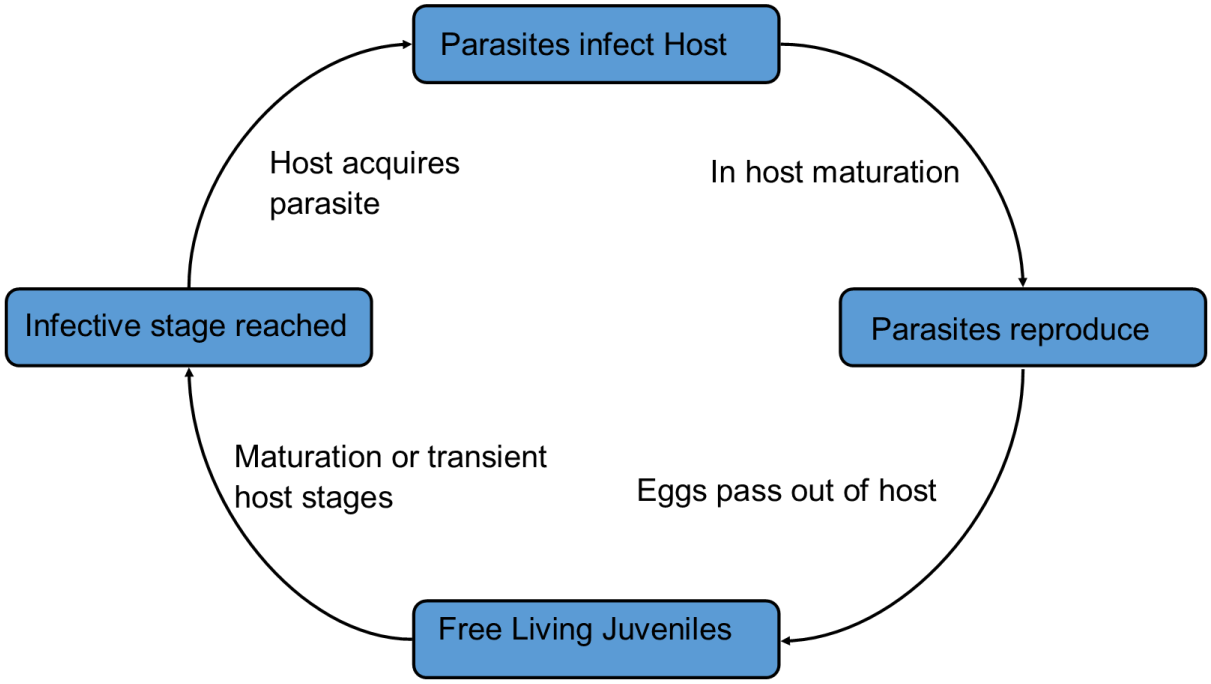


Figure 1-1: Basic life-cycle of a parasitic helminth.

4) Eggs either hatch and live external to a host as juvenile parasites or are picked up by a transient host.

5) Parasites mature to their infective stage infect a new host. (Return to 1)

A representative figure of this may be seen in figure 1-1. It may take many months for a juvenile parasite to infect a host and it may live for several years as an adult in the host after this [13, 94]. This life-cycle means that the spread of the infection is dependent not only on how many hosts are infected but also on their individual parasite loads and how many juvenile parasites are living within the external environment. Even if no hosts are currently infected, parasites living externally to the hosts can still infect a new host and re-start the infection [73].

In modelling macroparasite infections such as these, there is a major challenge in determining how the within-host dynamics occur and the rates at which different processes happen [68]. This is made even more difficult by the between host heterogeneity in all aspects such as pick up rates, carrying capacity, and immune response to parasites[37, 49]. Fortunately, research has been conducted previously into more general patterns of parasitic infection, which we may make use of to inform our model. Numerous researchers have discussed the tendency of macroparasites to have highly aggregated distributions, with a large degree of overdispersion [7, 18, 37, 56, 99]. This means that the variance in parasite burden exceeds the mean value across all hosts and the distribution of parasite burdens is highly skewed [93]. Initial parasite research often assumed that the distribution of parasites would be close to a Poisson distribution, which would have a mean equal to the variance. This has led to the aggregated distribution being referred to as overdispersed, meaning it has a

variance greater than expected. Often the negative binomial distribution has provided a good fit for these parasite distributions to accommodate the overdispersion [7, 5, 79]. With this being the case then a model that accounts for this may be of great use in helping determine the best course of action in treating these infections. Additionally, this evidence is used as a basis for some of the modelling techniques that we discuss throughout this thesis.

Macroparasites also provide a good motivating example for this research as the World Health Organisation has earmarked several macroparasitic diseases as important targets for control, such as Schistosomiasis and other soil-transmitted helminthiases [21, 78, 71] However control can create issues in and of itself. Macroparasite resistance to treatment is well documented [34, 96, 82]. Overtreatment combined with the difficulty in properly eradicating them has led to many parasite species showing increasingly high resistance to different treatments. The use of secondary hosts and the external environment as reservoirs, combined with lowly infected hosts often being asymptomatic and being neglected from treatments can further exacerbate the problem. This is a problem that has been observed a great deal in agricultural settings, such as widespread resistance to Benzimidazole in sheep and goats, [4, 100, 28, 82]. This tendency to develop resistance motivates the second half of the research we present in this thesis, as we seek to both model the dynamics of increasingly resistant populations, and consider how this may affect the efficacy of control strategies.

1.3 Previous Models of Macroparasites

Some of the most influential research in this area has been the work conducted by Anderson and May [7, 67]. They created a deterministic model for the size of the host population, H, and the total parasite population, P.

The model they derived considered a simple linear birth-death model for the host population, with hosts birth rate b and non-parasite related death rate μ_H. Hosts were also subject to a parasite related death rate due to parasites, calculated by assuming that each parasite caused an increase in the death rate of magnitude α. The average death rate due to parasites was then given by $\alpha H \sum_i i p_i$, where i was the number of parasites in a host and p_i was the probability of this being the case. The resulting function gave total host death due to parasites as $\alpha H \frac{P}{H}$.

The parasite population itself was subject to a birth/immigration rate which was dependent on the total number of parasites but also the host density, the resulting function was given by $\frac{\lambda P H}{H_0+H}$. This is given by the instantaneous birth rate λ per parasite, which leads to a total number of instantaneous births as $\lambda P(t)$. Anderson and May's model assumes that new parasites (larvae) wait in the external environment, not an intermediate host, before being picked up by a new definitive host. While they are in the external environment they are subject to death due to environmental factors and predation so only a proportion of them will survive to be picked up. This proportion depends on the density of hosts in the environment, relative to the other factors which may remove the larvae from the

parasite reservoir. From this, the larval dynamics may be modelled using

$$\frac{dL}{dt} = \lambda P - \mu_L L - \delta H L,$$

where μ_L is the rate at which larval parasites die due to external factors and H is the number of hosts. Using a time-scale justification we may assume that the larval population is at equilibrium. When this is the case the pick up rate per host, δL, is given by

$$\delta L = \frac{\lambda P}{\mu_L/\delta + H},$$

where $\mu_L/\delta = H_0$. Multiplying through by the number of hosts gives the final birth/immigration rate above.

Perhaps one of the most vital aspects of Anderson and May's work was the parasite death rate, aside from a basic within-host death rate, μ_P, there was also the parasite death which occurred as a result of the host death. When hosts died of causes unrelated to parasites at rate μ_H then the parasite death due to host death was calculated as $\mu_H H \sum_i i p_i = \mu_H P$, similar to the host death rate due to parasites. However when we consider parasite death due to host death caused by the parasites then the parasite death rate becomes $\alpha H \sum_i i^2 p_i = \alpha H \mathbb{E}[i^2]$. This death rate introduced the second-order moment of the parasite distribution throughout the host population into the model. This may be seen in the model equation of their model which are:

$$\begin{aligned}
\frac{dH}{dt} &= (b - \mu_H)H - \alpha P \\
\frac{dP}{dt} &= \frac{\lambda P(t) H(t)}{h_0 + H(t)} - (\mu_H + \mu_P)P(t) - \alpha H \mathbb{E}[i^2].
\end{aligned}$$

Anderson and May's method to close this system and calculate $\mathbb{E}[i^2]$ was to use a distribution that was believed to model the true parasite distribution throughout the host population and, by fixing parameters associated to this distribution, derive a function for the second-order moment based on the mean parasite population. One of the key results of the work involved using a non-random distribution, meaning that the location of parasites is not independent of the other parasites, often occurring in clumps. Anderson and May specifically used the negative binomial distribution and found that doing so stabilised the model, and allowed the system to settle to equilibrium. In contrast when the parasites were assumed to be distributed between hosts according to a Poisson distribution the model exhibited cyclic changes in the states. Although Anderson and May considered the effects of numerous modifications to the model it is this property of using empirical studies of parasites to inform the model itself which has made their research so influential.

Building off the work of Anderson and May numerous other researchers have gone on to try and incorporate the non-linear effects of parasite populations. Grenfell *et al* [37], and Kretzschmar and Adler [56] both targeted their research on aggregated distributions. In the case of Grenfell et al,

this was in regard to how host age affected the potential aggregation of parasite distributions. Their model considered a single host, with the infection described by four variables, l, m, e, and i. The variables l, m, e denoted the number of larvae, mature parasite burden, and egg load associated with the host, and the immune response of the host was modelled by i. Grenfell *et al* assumed that the host exposure and death rates were dependent on the host's age. As a host aged and was exposed to more parasites the immune response increased and with it the host's ability to fight the parasite infection. With this, they formulated a model for the joint distribution of the four variables. Although they ultimately concluded that allowing for additional sources of variability was important, the method by which they modelled provided them with the ability to examine the relationship between variables.

The work of Kretzschmar and Adler [56] considered a different approach, by considering the variance to mean ratio of parasite burdens as a variable itself. This altered the model such that the function used to approximate the second-order moment no longer requires fixed parameters of the estimated distribution. Using these models they were able to consider results when different assumptions were made about the underlying distributions of parasites, especially for aggregated distributions. The model that they derived is given by:

$$\begin{aligned}
\frac{dN}{dt} &= N(\beta - \mu - \alpha x) \\
\frac{dx}{dt} &= x\left(\frac{\kappa P}{c+N} + \rho - (\beta + \sigma) - \alpha\pi\right) \\
\frac{d\pi}{dt} &= \beta x + 2\rho - (\pi - 1)\left(\frac{\kappa N}{c+N} + \sigma + \alpha - \rho + \delta\alpha(\pi - 1)\right).
\end{aligned}$$

Here the variables N, x and π denote, respectively, the number of hosts, the mean parasite burden per hosts ($x = \frac{P}{N}$), and the variance to mean ratio. The rate parameters in this model are given by

- β is the birth rate of new hosts.
- μ denotes the death rate of hosts due to non-parasite related causes.
- α is the increase in host death rate caused by each parasite.
- The function $x\frac{\kappa N}{c+N}$ denotes the immigration rate of new parasites into hosts. This is found by the same method used in the Anderson and May model described previously.
- The model also includes the possibility that parasites may reproduce within the host, and the rate at which this occurs is given by ρ.
- The final parameter, δ, is a distribution parameter introduced into the model to close the system, based on assumptions about the underlying distribution. When $\delta = 1$ it is assumed to be negative binomial, and when $\delta = 0$ it is assumed that the parasites are distributed according to a Neyman type A distribution, which is essentially a modified Poisson distribution which corresponds to a random distribution [27]. Specifically, a Neyman type A distribution is a Poisson distribution with a parameter given by a constant $\phi > 0$ multiplied by a random

variable $X \sim Pois(\lambda)$ [65].

One of the key findings of this work was that altering the assumed distribution could have distinct effects on the outcome of the model, changing the equilibria and even the stability. This demonstrates the need for well informed choices around the underlying distribution of aggregated populations, such as parasites, in models such as these. That is to say, if a distribution is chosen well it may edge the dynamics into a more realistic representation of the true system. This is a property Isham [49] addressed when considering a deterministic approximation to a stochastic system, which is discussed in greater length in chapter 2.

Some of the other models which have proven particularly influential to this research are those of Roberts and Grenfell [84, 85]. This is because these models were formulated on the basis of a managed agricultural population, just as the models formulated in this thesis are. One of the primary inclusions in both their and our models, which results from this assumption, is that the host population size is constant. This is very different from the other models discussed here, chiefly as this means that hosts do not die due to parasite infection. The base model that they derived is:

$$\begin{aligned}\frac{dL}{dt} &= -(\rho+\beta)L + q\lambda(r)P \\ \frac{dP}{dt} &= \beta p(r)L - \mu(r)P \\ \frac{dr}{dt} &= \beta L - \sigma r,\end{aligned}$$

where L denotes the mean density of a free-living larval population in the area associated with the host, P is the mean intensity of the parasite infection in each host, and r is the level of acquired immunity. In this model the parameters which determine the dynamics are:

- β denotes the rate at which larval parasites are ingested by a host.
- The probability of ingested larvae maturing is given by $p(r)$, a monotonic non-increasing function of the immunity level.
- Mature parasites die off within the host at a rate, $\mu(r)$, which is also dependent on the immune level. In this case, the rate is a monotonic non-decreasing function of the immunity level.
- $\lambda(r)$ is the rate of egg production of the mature parasites.
- The probability q denotes the probability that an egg hatches and reaches the larval stage.
- Larval parasites suffer a death rate of ρ while they are living externally to the host.
- σ is the rate at which immunological memory fades.

With this model, they found that one of the most interesting properties was that before reaching equilibrium there was typically a peak in parasite numbers, which then dropped to a lower value at equilibrium. For a managed population such a feature has implications on control as, when monitor-

ing an infection, control may appear necessary when this peak occurs, but the parasite population would actually stabilise and drop naturally.

Expanding on this work, Roberts and Grenfell conducted research into how seasonal fluctuations could affect parasite populations [85]. Given the nature of the parasite life cycle, free-living parasites are particularly susceptible to changes in weather that occur periodically throughout the year. To adapt for seasonality they used forcing functions to alter the larval transmission rates. In their conclusions they found that seasonal fluctuations could lead to increases in the parasite and immunity levels.

While biological models are often formulated as deterministic models, the natural stochasticity of populations is certainly important, as has been discussed extensively by Nisbet and Gurney [76]. In particular, Marion et al [61, 62] used the model formulated by Roberts and Grenfell [84] and modified it using a variety of methods to allow for variation in the different processes, such as parasite birth, death and pick up rates. In doing this they described how, although deterministic models may be useful for describing general behaviour, the randomness that isn't accounted for in these models could lead to substantial fluctuation around the endemic state. Perhaps some of the most important research into stochastic modelling of macroparasite populations with regards to this thesis is that of Isham, and Isham and Herbert [49, 47]. The models that they worked with considered the stochastic basis of a macroparasite model and they used this to model the rate of change for the whole distribution of parasite burdens in hosts. They then discussed how moment closure approximations could be used to improve estimation in situations where the full distribution could not be determined. We will return to this model approach in chapter 2.

1.4 Metapopulation Modelling

Another vital field of research for this project is that of metapopulation models. According to Hanski, metapopulation models traditionally fall into three categories [41, 42, 40].

The first is two population models. Although they can involve more than two populations, these are models that examine the local dynamics of the populations where, although separate, individuals may migrate between populations. These may take a form such as:

$$\frac{dN_1}{dt} = r_1 N_1(1 - \frac{N_1}{K_1}) + \lambda_1 N_2 - \lambda_2 N_1$$
$$\frac{dN_2}{dt} = r_2 N_2(1 - \frac{N_2}{K_2}) + \lambda_2 N_1 - \lambda_1 N_2.$$

Here N_1 and N_2 denote the two populations within the two environments. Without immigration the populations dynamics are both described by a logistic growth equation with growth rates r_1 and r_2 and carrying capacities K_1 and K_2 [74]. However, there is a link between the two populations

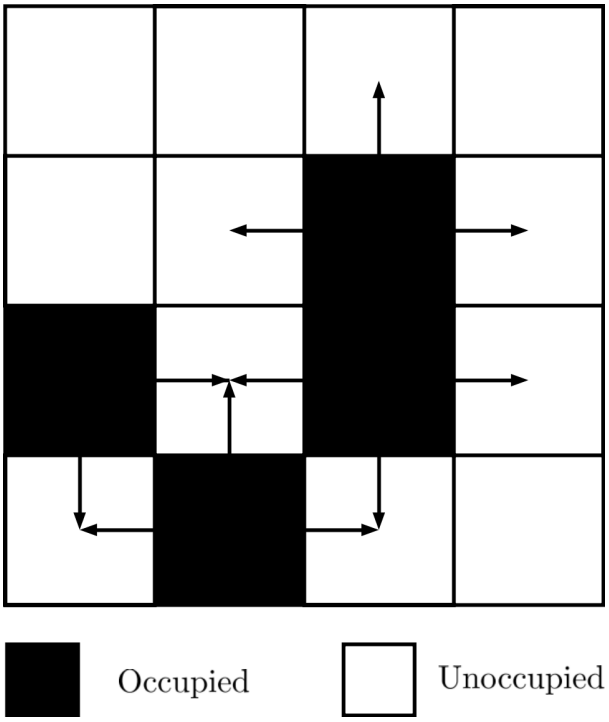


Figure 1-2: Basis of a lattice model with migration from occupied cells to adjacent cells

allowing individuals from each population to move between them at rates λ_1, from population two to population one, and λ_2, in the other direction. The advantage of these models is that the populations are able to vary according to their individual subpopulation dynamics and behave differently to one another depending on the environmental factors which uniquely affect the separate patches they inhabit. The disadvantage comes when this is expanded to large numbers of populations, it is difficult to assign individual parameters to all populations and results in very large systems of equations. It also introduces more sources for error as these parameters must be determined for every patch.

The two other categories of models that Hanski describes are lattice models and Levin's model. Levin's model is a model that considers a very large number of patches which are either occupied or not [31], similar to a standard SIS infection model. Local dynamics are ignored in favour of simply deriving a model to describe the percentage of patches that are occupied. Levin's model makes certain simplifying assumptions, chiefly that migration from a population is equally likely to spread to any of the other populations, essentially ignoring any spatial dynamics or environmental variation. Lattice models build on the principles of the Levin model but each patch is located on a lattice. Migration into a cell occurs from neighbours at a rate proportional to the number of neighbours occupied [92]. The basis of a simple lattice model may be seen in figure 1-2.

For macroparasitic infection models transmission dynamics are not as strongly related to a spatial factor as the patches, or hosts, move about and the transmission is indirect, first going through an external phase. As a result, a lattice model may not be necessary. However, given the secondary

focus of this research on the use of Anthelmintics (sometimes known as Antihelmintics) to control parasite infections and the effect of anthelmintic resistance on population dynamics, the local dynamics are likely to be of great importance. This returns us to the idea of the two population models. As discussed, this can lead to large systems of equations and parameter values so we instead look to a different method. Keeling published research that informs the work here greatly [52] and is discussed in the methodology presented in chapter 3.

1.5 Summary

In this chapter, we have discussed previous research into macroparasites, including an examination of the biological properties of macroparasites that inform mathematical models and discuss why they are of such interest as a control problem. While we have also discussed some of the mathematical research in this area the study of macroparasitic diseases is a wide-ranging field. In the following chapters we aim to further investigate certain properties of macroparasites and the relation to control. Further background into topics that inform this are discussed at the start of each chapter, including the method of model formulation, analysis of models, and optimal control.

In the next chapter we discuss in far greater detail the work of Isham [49] and use it as a basis by which we consider the potential implications of control on a macroparasitic model. Chapter 3 details the formulation of our linked population model, presenting two formulations using moment closures for comparative purposes. In chapter 4 we discuss how control is implemented, beginning from the basis of the model. Following this, we formulate how a control regime influences the dynamics of the models to examine how alternative formulations of the model change the most effective control. Parasite resistance to treatment is introduced to the model in Chapter 5, through a combination of genetic and competition models. Finally, chapter 6 and 7 reformulate the optimal control problems using the resistance models to assess how the properties of these models alter the optimal control when compared with the models from chapter 3. The appendices detail all of the Matlab codes that were written to produce the numerical simulations presented in the thesis, directing the reader to associated github files.

Chapter 2

Preliminary Models

Before looking at how best to control macroparasitic infections, we first looked into different models of macroparasite dynamics. This allowed us to determine the most appropriate models for the specific problems of interest, namely the effect of variance in parasite load on control problems. Although these models included both fully stochastic models and deterministic models, this chapter is primarily based on work previously done by Herbert and Isham [49, 47]. These were based on fully stochastic models that enabled them to examine the full distribution of macroparasites. Through moment closure approximations they could approximate the mean and variance of parasite burdens using ordinary differential equations (ODEs). The hope in doing so was that for more complex systems in which the full distribution was indeterminable approximations of the moments of the parasite burden distribution could be formulated instead. The aims of this chapter are

- To detail the model formulated by Isham [49] and use it to show how moment closure approximations may be used to give reasonable deterministic approximations to fully stochastic systems;
- To introduce a basic form of control to Isham's model as a preliminary investigation into how the implementation of a control measure on an overdispersed set of sub-populations will affect the control;
- To discuss the difficulty with this model when advancing further and introduce the stochastic basis of the alternative models that are derived in a later chapter.

We first provide a brief overview of a model formulated by Isham [49]. This model provided a useful starting point for work developed later in this thesis. The model formulation by Isham was based on stochastic processes which govern the size of a parasite population within a host. If we denote the parasite population size in a host at time t, by $M(t)$ then the crucial processes and rates which the model incorporates are as follows:

- Hosts pick up clumps of parasites from a external source, e.g a pasture. A clump is a group of parasites of variable size C, which is a non-negative integer-valued random variable. The

clump size distribution has probability generating function $h(z) = \sum_{c=0}^{\infty} h_c z^c$ [93]. It is assumed that clumps are picked up at random times. The difference between each of these successive time points is given by $\tau_j = t_{j+1} - t_j$.These τ_j are distributed according to an exponential distribution with rate parameter ϕ, that is $\tau_j \sim \exp(\phi)$ for all j.

- Parasite death occurs within the host at a rate of μ_M per parasite.
- Host death due to non-parasite related causes, occurs at a rate of μ_H.
- The presence of parasites causes an increase of size α per parasite in the host death rate.

Using these assumptions Isham formulated an equation for the rate of change of $\mathbb{P}_M(t)$, at time t, where

$$\mathbb{P}_M(t) = \text{probability a host is alive at time } t \text{ and is infected with } M \text{ parasites.}$$

This probability master equation takes the form:

$$\begin{aligned}\frac{d\mathbb{P}_M(t)}{dt} = &-(\mu_H + \alpha M + (1-h_0)\psi + \mu_M M)\mathbb{P}_M(t) + \mu_M(M+1)\mathbb{P}_{M+1}(t) \\ &+ \psi \sum_{c=1}^{M} \mathbb{P}_{M-c}(t) h_c \end{aligned} \quad , \tag{2.1}$$

where $1 - h_0$ denotes the probability that at least one new parasite was picked up. Using the master equation, Isham derives a differential equation for the probability generating function for the number of parasites in a host at time t given that it survives to this time. To do this, a generating function for the number of parasites within a host and its survival to time t is defined as:

$$P(t; z) = \sum_{M=0}^{\infty} \mathbb{P}_M(t) z^M, \tag{2.2}$$

to give a time-dependent probability generating function. Using this definition with (2.1) the partial derivative with respect to time of this generating function is given by:

$$\frac{\partial P(t;z)}{\partial t} = -(\mu_H(t) - \phi(h(z) - 1))P(t;z) - ((\alpha + \mu_M)z - \mu_M)\frac{\partial P(t;z)}{\partial z}. \tag{2.3}$$

Isham then defines $S(t)$ by setting it equal to $P(t;1)$ which is the probability that a host survives to time t and $\mathbb{E}(M(t)) = \left.\frac{1}{S(t)}\frac{\partial P(t;z)}{\partial z}\right|_{z=1} = \left.\frac{\partial Q(t;z)}{\partial z}\right|_{z=1}$ is the expected number of parasite within a host given that it survives to time t. The conditional generating function $Q(t;z) = \frac{P(t;z)}{S(t)}$ is defined using Bayes' theorem. Setting $S(t) = P(t;1)$ in (2.3) leads to

$$\frac{dS(t)}{dt} = -\mu_H(t)S(t) - \alpha\mathbb{E}(M(t))S(t). \tag{2.4}$$

Isham determines, by the quotient rule, that the conditional generating function obeys:

$$\frac{\partial Q(t;z)}{\partial t} = (\phi(h(z)-1) + \alpha\mathbb{E}(M(t)))Q(t;z) - ((\alpha+\mu_M)z - \mu_M)\frac{\partial Q(t;z)}{\partial z}. \tag{2.5}$$

Once solved $Q(t;z)$ can be used to give the mean, variance, and higher order moments of the distribution of parasites. In this case (2.5) could be analytically solved but the methodology that Isham gives to determine the mean and variance when this is not the case is of particular interest in this thesis, as it is a method which we employ in the formulation of our models.

2.1 Moment Closure Approximation

One of the main parts of Isham's methodology that is of interest throughout this thesis is their use of moment closure approximation to create closed differential equation systems that model the mean and the variance of the parasite populations [52, 48, 7]. The negative binomial distribution has been shown to be important in macroparasitic infections, which makes its use in the approximation of the mean and variance of the parasite burdens particularly interesting [5]. In order to demonstrate this moment closure approximation method, Isham considers a model which has the distribution of the clumps picked up also given by a negative binomial distribution. Evaluating the partial derivatives with respect to z of $\dfrac{\partial Q(t;z)}{\partial t}$ at $z=1$ gives differential equations that describe the rate of change of the mean, $\mathbb{E}(M(t))$, and the variance, $V(t)$,of the parasite load with respect to time. These are as follows:

$$\begin{aligned}
\frac{d\mathbb{E}(M(t))}{dt} &= \left.\frac{\partial^2 Q}{\partial t \partial z}\right|_{z=1} &&= \phi h'(1) - \alpha V(t) - \mu_M \mathbb{E}(M(t)) \\
\frac{dV}{dt} &= \left.\frac{\partial^3 Q}{\partial t \partial z^2}\right|_{z=1} &&= \phi(h''(1)+h'(1)) + \mu_M\mathbb{E}(M(t)) - 2\mu_M V(t) \\
& && \quad + \alpha(3\mathbb{E}(M(t))V(t) + \mathbb{E}^3(M(t)) - \mathbb{E}(M(t)^3)).
\end{aligned} \tag{2.6}$$

In (2.6) the term $\mathbb{E}(M(t)^3)$ refers to the third-order moment. This is where the moment closure approximation is applied with the relationship between the third-order moment and the lower order moments chosen based on a distribution that is thought to most accurately represent the true distribution. In the case of a moment closure approximation based on the negative binomial distribution this relationship, Isham gives

$$\mathbb{E}(M(t)^3) = 3\mathbb{E}(M(t))V(t) + \mathbb{E}^3(M(t)), \tag{2.7}$$

which upon substitution into the system in (2.6) gives the system

$$\begin{aligned}
\frac{d\mathbb{E}(M(t))}{dt} &= \phi h'(1) - \alpha V(t) - \mu_M \mathbb{E}(M(t)) \\
\frac{dV(t)}{dt} &= \phi(h''(1)+h'(1)) + \mu_M\mathbb{E}(M(t)) - 2\mu_M V(t).
\end{aligned} \tag{2.8}$$

Parameter/Distribution	Value	Unit	Description
$h(z)$	$NB(p,r)$	no units	Distribution of clump sizes (Negative Binomial)
p	0.9474	no units	Parameter of negative binomial distribution, $p \in (0,1)$
r	0.5	no units	Parameter of negative binomial distribution, $r > 0$
ϕ	52	years^{-1}	Average rate of clump pick ups
μ_M	10	years^{-1}	Average rate of parasite death
α	0.02	years^{-1}parasite^{-1}	Rate increase in host death caused by a single parasite
μ_H	0	years^{-1}	Average rate at which hosts die
u	20	years^{-1}	Increase in average parasite death rate caused by treatment

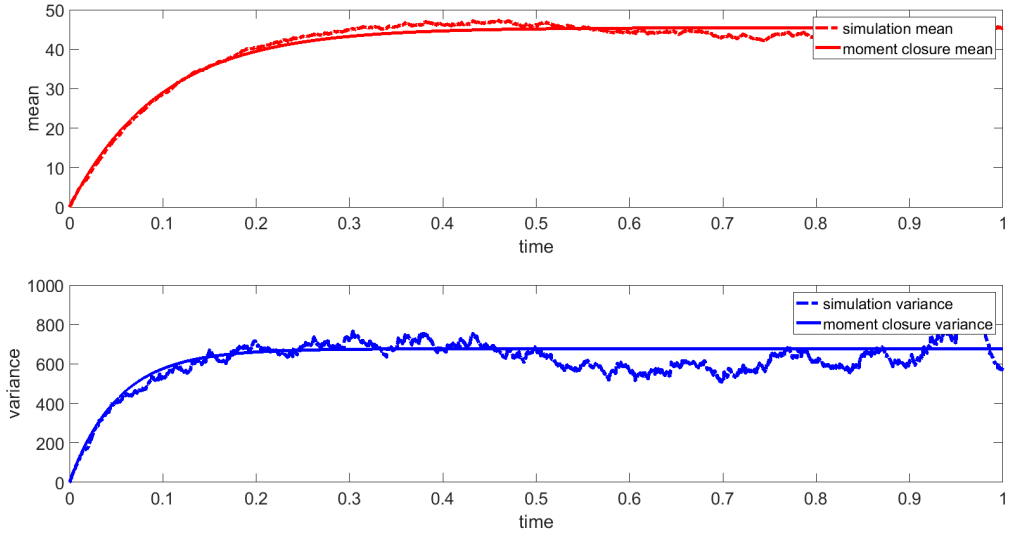
Table 2.1: Parameter values for Isham model example and three treatment models

2.1.1 Model Examples

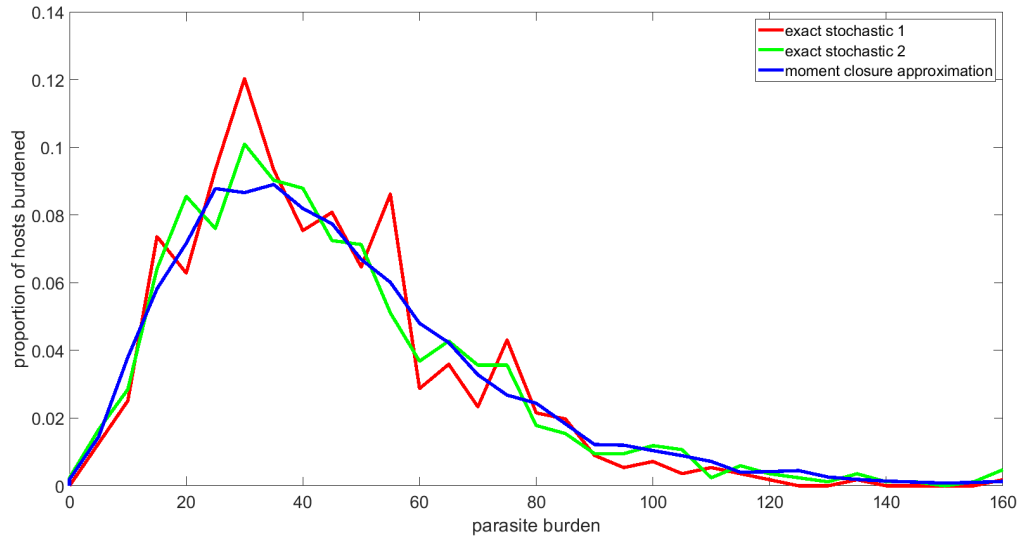
With the approximated model, (2.8), a crucial part of Isham's work involved the comparison of the approximation with simulations of the full stochastic system. In doing so they were able to establish whether the method of approximation provided useful results. Before the model may be simulated either stochastically or using the moment closure approximated system, the parameter values must be set. Table 2.1 shows the parameter values used for the rates of the processes and distributions used in the model. The parameter u is used in later sections as a control parameter resulting from treatment and was chosen to be of the same order of magnitude as the parasite death rate. It is then assumed that clumps are picked up approximately once a week, parasites live within the host for a little over a month, and host death without parasites in that time frame is assumed to be negligible so μ_H may be set to zero. The parameters of the distribution of the clumps are chosen such that each clump has a mean size of 50 and a variance of approximately 210, which would be appropriate for a negative binomial distribution where the variance exceeds the mean. Host death due to parasites is low when parasite numbers are low and the majority of hosts are alive at the end of simulations to facilitate interpretation of the results.

Figure 2-1 sub-figure (a) shows a comparison between the mean and variance of the simulated model taken over 100 repetitions compared with the mean and variance given by the moment closure approximated mean and variance. Here the simulations are performed using the Gillespie algorithm [35] combined with the original rates and processes which Isham defined. Examining this figure, the approximated mean estimates the simulated mean quite well. The approximated variance appears to estimate the simulated variance less well. This is not surprising, especially given the limited number of repetitions of the simulation. Although more simulations could have been performed this was kept limited as the work of Isham has already looked in greater detail at this. When considering the error between the simulated variance and approximated variance relative to its magnitude, it performs far

better than it appears. Figure 2-2 shows the error in the simulation mean and variance relative to



(a)



(b)

Figure 2-1: (a) The mean and variance of 100 stochastic simulations of Isham's model and the mean and variance given by a moment closed approximation of the mean and variance (b) Two examples of the distribution of parasites from stochastic simulations and the approximate distribution using the mean and variance from the moment closure. Note that we represent the discrete burdens using continuous lines for the ease of reading the graph. Model parameters given in table 2.1.

the size of the mean and variance predicted by the moment closure approximated version of Isham's

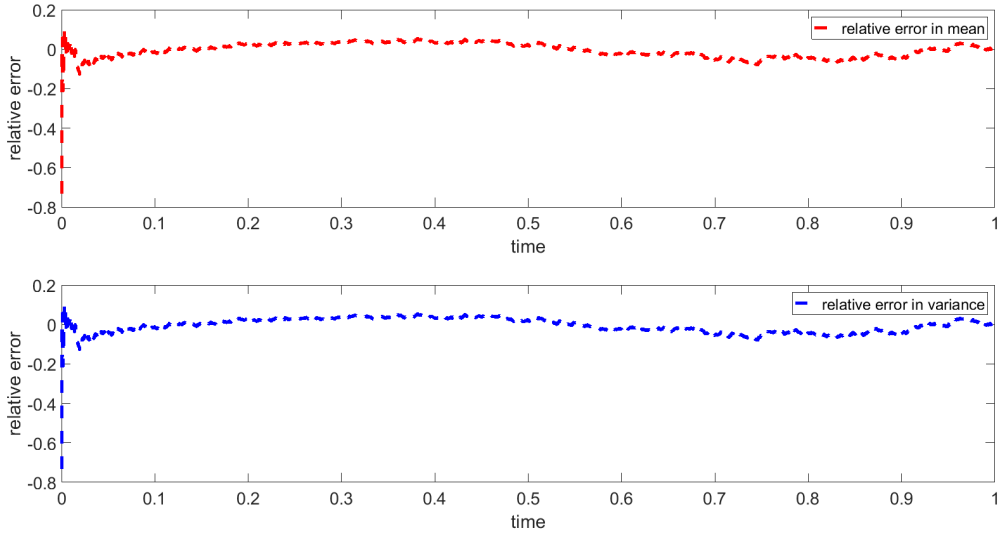


Figure 2-2: Error in simulations relative to the size of the mean and variance from the moment closure approximation from Isham's model. Parameter values as given in table 2.1.

model. These errors are calculated by using

$$e_M = \frac{\mathbb{E}[M(t)] - \frac{1}{N}\sum_{j=1}^{N} M_j(t)}{\frac{1}{N}\sum_{j=1}^{N} M_j(t)}, \qquad e_V = \frac{V(t) - \left(\frac{1}{N-1}\sum_{j=1}^{N}(M_j(t) - \frac{1}{N}\sum_{k=1}^{N} M_k(t))^2\right)}{\frac{1}{N}\sum_{j=1}^{N}(M_j(t) - \frac{1}{N}\sum_{k=1}^{N} M_k(t))}.$$

Figure 2-1, sub-figure (b), looks at how the moment closure approximated mean and variance may be used to try and look at an approximation of the whole distribution against what the actual distribution is in the simulation at different time points. If it is assumed that the true distribution of the parasite loads per host is given by a distribution of the same type as was chosen for the basis of the moment closure approximation, in this case the negative binomial distribution, we may try to determine an approximation of the full parasite distribution. The approximated mean and variance in parasite load from the ODE model are used to determine approximate parameters for the distribution that is believed to fit the full simulated data of the parasite burdens. For the negative binomial distribution these are [93]:

$$\begin{aligned} p &= \frac{\mathbb{E}[M(t)]}{V(t)} \\ r &= \frac{\mathbb{E}^2[M(t)]}{V(t) - \mathbb{E}[M(t)]}. \end{aligned} \tag{2.9}$$

With these parameters, the negative binomial distribution may be plotted and compared with the true distribution of parasites across the simulations. Comparing the distributions of parasites as given by the simulations with a negative binomial distribution with the parameters found in (2.9) from the approximated mean and variance shows that the approximation and the simulation appears qualitatively similar, although there is inevitably some deviation between the estimates. Moreover,

the shape of the distribution, figure 2-1 sub-figure (b), clearly shows the overdispersion that is of interest in this research.

On the basis of this model by Isham the main question of this research may begin to be addressed; how does including the variance, particularly on an overdispersed distribution affect the treatment strategy?

In the next section, we use the model that Isham formulated to examine how distributions may change when different, individual-based treatments are applied. To do this, we consider a combination of the stochastic simulation model and the moment closed model.

2.2 Application of Treatment

In order to optimise treatment, it is first important to establish how it is applied and what effect this has on the model. This section takes the models that Isham [49] formulated and begins to examine how these may be advanced to consider control strategies. A simple control is considered in which host treatment causes the parasite population within that host to suffer an increased linear death rate μ_M for the duration of the treatment. The formulation of these control dynamics is discussed at greater length in chapter 4 in the context of the models we derive in chapter 3. As we have shown in figure 2-1, the level of infection of individual hosts is extremely varied. As a result of this, treatments on different individuals will affect the overall distribution in very different ways. If a lowly infected host is treated the total reduction in parasites is far more limited than if a host with a very high number of parasites is treated. This is a property which is of especially great importance in models in which the parasite populations in the individual hosts are linked to one another. Although the hosts in the Isham model are not linked it is still a useful model to begin to consider the effects of different individual-based treatment strategies. To begin to investigate the specifics of this, we incorporate the following three treatment strategies into a model:

- Treat all hosts at set times.
- Treat hosts as soon as they become burdened above a certain threshold.
- Treat hosts at set times if they are considered overburdened at the start of the treatment period.

2.2.1 Treatment at Set Times

The first strategy it is assumed that treatment occurs at set times and is applied to every host. This is a simple adaptation to make to Isham's model. As all hosts are treated simultaneously and the same effect is seen on all of them, this may be modelled by an additional death rate, $u(t)$, on the

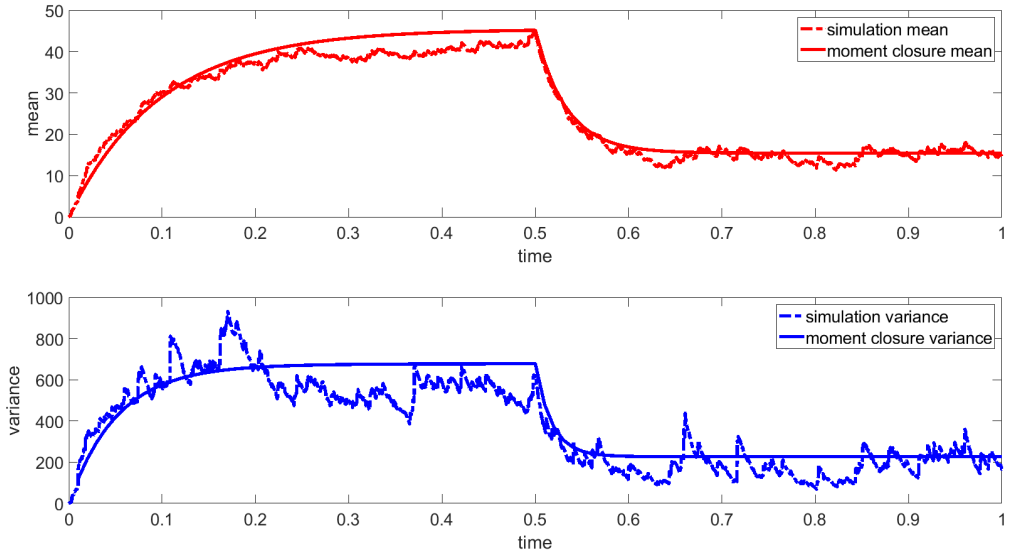
parasites during treatment, with the treatment death rate modelled by a linear death rate

$$u(t) = \begin{cases} 0 & \text{if treatment is off} \\ u_{max} & \text{if treatment is on.} \end{cases}$$

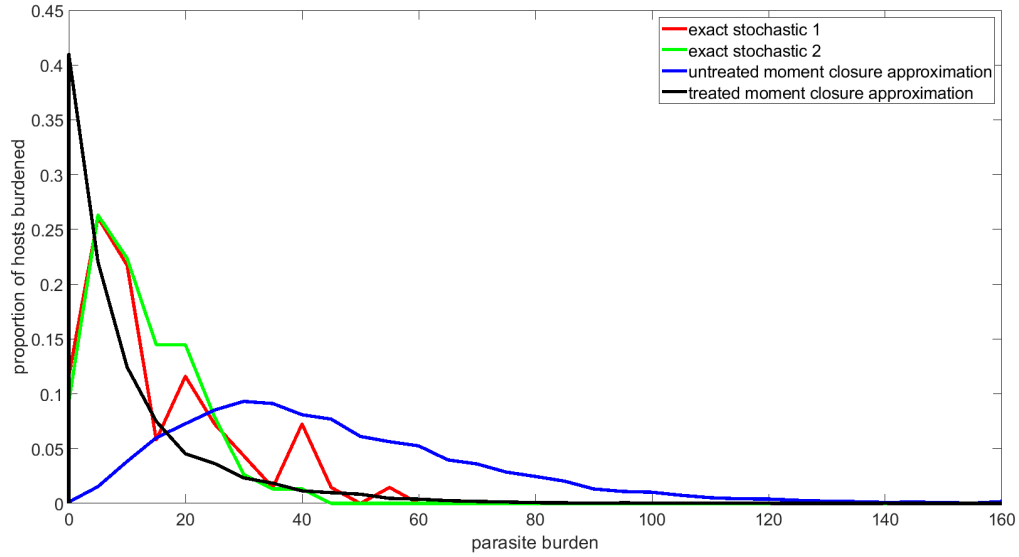
Analysis of the model when treatment follows exactly as for the basic model. As a result of this, the model will be given by

$$\begin{aligned} \frac{d\mathbb{E}(M(t))}{dt} &= \phi h'(1) - \alpha V(t) - (\mu_M + u(t))\mathbb{E}(M(t)) \\ \frac{dV(t)}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M + u(t))\mathbb{E}(M(t)) - 2(\mu_M + u(t))V(t), \end{aligned} \tag{2.10}$$

with $u(t)$ taking the value that corresponds to whether treatment is being applied or not. As the treatment occurs at pre-set times it will be known whether treatment is being applied to a host or not and what value u will take without the need for full simulation. The behaviour of the model at a specific time can be determined from the approximated model and whether or not treatment is active. The codes for this may be seen in Appendix A.1. Figure 2-3 shows the effects of switching on control at time $t = 0.5$ and keeping it on for the remainder of the simulation. Much like for the untreated approximation, the moment closure approximation continues to work well and give good approximations of the simulated system for the treated portion of the simulations, as would be expected. The distribution shifts quite significantly so that far more hosts are infected with 10 or fewer parasites and far fewer hosts are infected with more than 10 parasites in comparison to the untreated distribution. This is demonstrated in sub-figure 2-3 (b) which shows a comparison of the untreated moment closure approximation with the treated moment closure approximation and examples of the simulations actual distribution at two separate time points during treatment.



(a)



(b)

Figure 2-3: (a) The mean and variance of 100 stochastic simulations of Isham's model compared with the mean and variance given by a moment closed approximation of the mean and variance, when control is applied to both models between $t_1 = 0.5$ and $t_2 = 1$. (b) Two examples of the distribution of parasites from different time points of the exact stochastic simulations and the approximate distribution using the mean and variance from the moment closure approximation before and after treatment. Model parameter given in table 2.1.

2.2.2 Parasite Burden Dependent Treatment

Rather than treating all hosts at a set time, hosts are now treated only when they meet a pre-defined criteria. In this particular example, we assume that for each host:

- Hosts start being treated when they have a parasite burden exceeding some critical level, C.
- Treatment is sustained for a fixed time period, T.

This gives a control function for an individual host as

$$u(t) = \begin{cases} u_{max} & \text{if } M(\tau) \geq C \text{ for some } t - T < \tau < t \\ 0 & \text{otherwise,} \end{cases}$$

which means that if the current time is t if the parasite burden exceeded the set value C at any point in the previous T units the treatment would have begun and would still be ongoing. With this treatment function if the parasite burden of a host is above the level C after the treatment has already been applied for some time the treatment period is extended.

This model is analytically challenging because host burden changes continuously over time and not every host will undergo treatment at once. Under these conditions for treatment we may rewrite equation (2.1) as

$$\begin{aligned} \frac{d\mathbb{P}_M(t)}{dt} = &-(\mu_H(t) + \alpha M + (1-h_0)\psi + (\mu_M + u\mathbb{1}(M(t-\tau) \geq C)M)\mathbb{P}_M(t) \\ &+ (\mu_M + u\mathbb{1}(M(t-\tau) \geq C))(M+1)\mathbb{P}_{M+1}(t) + \psi \sum_{c=1}^{M} \mathbb{P}_{M-c}(t)h_c \end{aligned} \tag{2.11}$$

where $0 \leq \tau \leq T$.

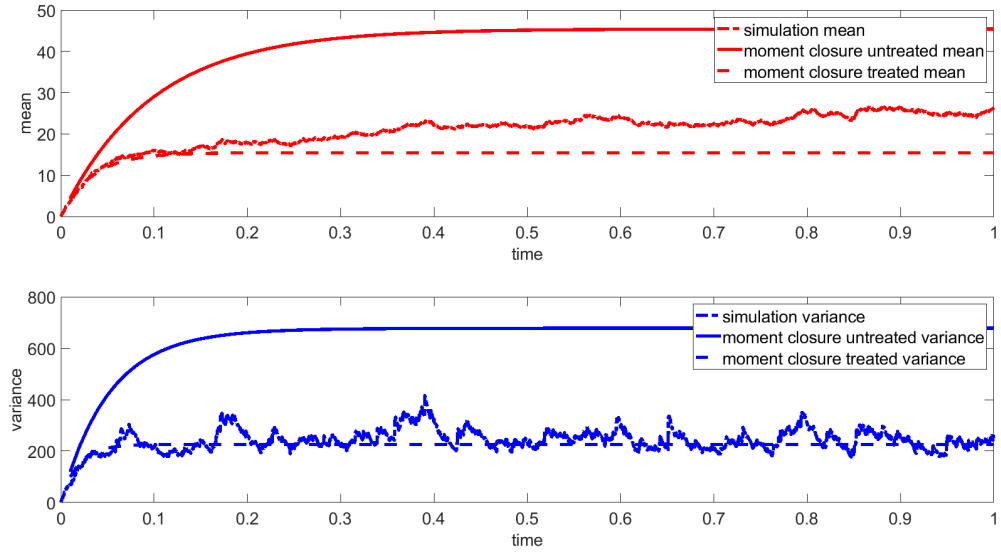
The system can be simulated without great challenge but analytical results around the impact of treatment are indeterminable. A comparison between the stochastic simulations for this scenario with the moment closure approximations is shown in (2.8), both with the standard death rate μ_M and the treated death rate $\mu_M + u$, is shown in figure 2-4 where hosts are treated when their burden exceeds $C = 50$ and a single treatment period last for a period of 0.05. These comparisons provide useful insights. If we look at the simulation mean then this lies between the fully treated and untreated means given by the moment closure model, as would be expected. The simulation variance however is, qualitatively, significantly closer to the variance given by the fully treated moment closure model, than to the entirely untreated variance. Based on the distribution in the absence of treatment, seen in figure 2-1 sub-figure (b), it would not be expected for a majority of hosts to be being treated at this time. As such these simulations demonstrate that individual based treatments can have a more substantial effect on the variance in parasite loads than on the mean parasite burden.

This highlights one of the motivators for this thesis; what happens if the control is intended to lower the variance as a priority, rather than just the mean?

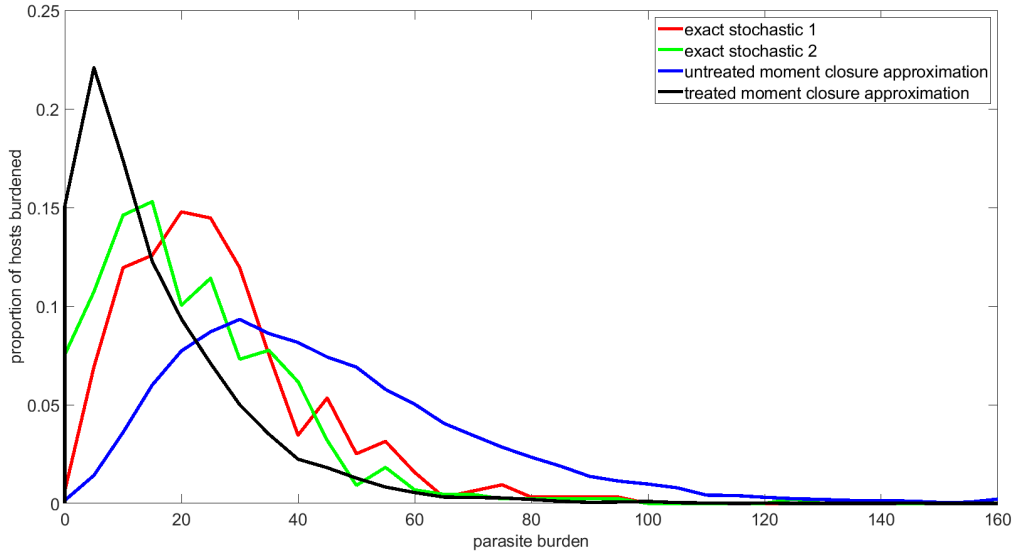
The effect of the treatment on the distributions can be seen in Figure 2-4(b). The distributions of the treated simulations have fewer hosts infected by very high burdens, that is above 40 parasites, than the entirely untreated approximate distribution. However, there are far fewer hosts infected by lower numbers of parasites (between 0 and 15) when the burden dependent treatment is applied than we see in the fully treated simulations (figure 2-3). In the simulations here, most hosts are now infected with burdens between 15 and 40 parasites. This is certainly due to the fact that only hosts that became burdened above the trigger level, in this case 50 parasites, will experience treatment so the shift does not act on the whole population so the distribution of parasites is now far less overdispersed.

2.2.3 Burden Dependent Treatment at Fixed Times

To reflect a more realistic scenario, as parasite burdens cannot be monitored constantly, we imposed the condition that treatment could only be switched on at set times and was applied only to hosts that were considered overburdened at that time. In modelling a treatment this way the hope is that there may be potential to use a moment closed model to approximate the simulations. This is discussed in greater detail in chapter 8, with the focus here being on the effect seen in the stochastic simulations. Figure 2-5 shows a comparison between the treated and untreated moment closure approximations of Isham's model and simulations of this level-dependent treatment switched on and off at discrete time points. For the simulations, the variance lies approximately between the moment closure approximation variances for the fully treated (section 2.2.1) and untreated models, as would be expected, but the mean of the simulations is much closer to the mean of the untreated moment closure approximation. It is not clear, however, precisely how this relates to the distribution of parasites through the host populations.

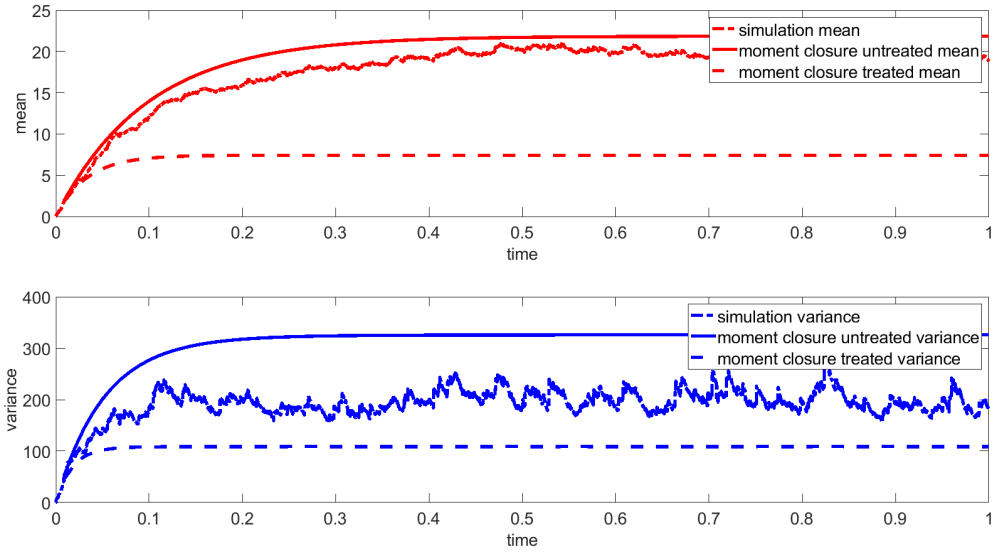


(a)

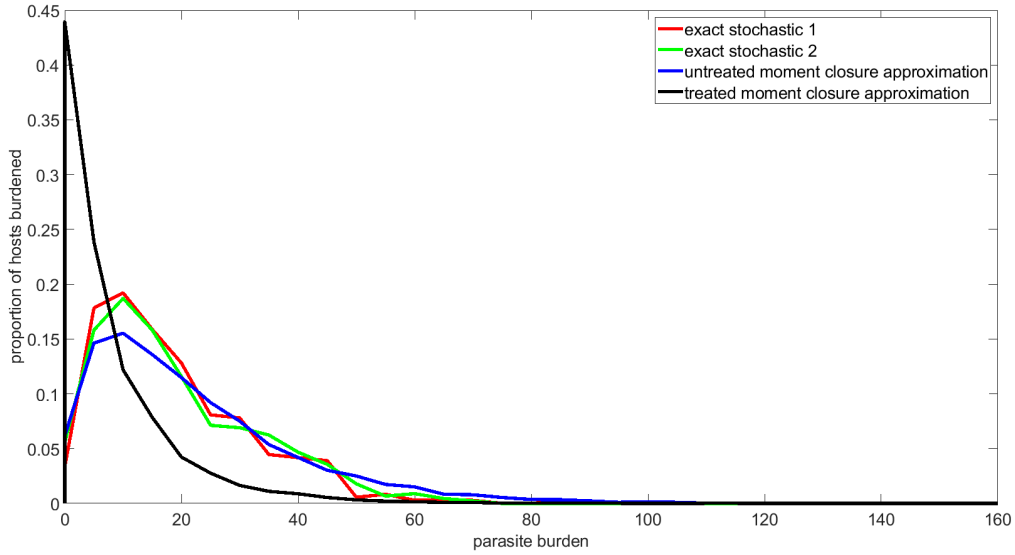


(b)

Figure 2-4: (a) The mean and variance of 500 stochastic simulations of Isham's model when a host is treated when its parasite burden exceeds 50 and the mean and variance given by a moment closed approximation of the mean and variance in the absence of treatment (b) Two examples of the distribution of parasites from stochastic simulations and the approximate distribution using the mean and variance from the moment closure approximations with and without treatment. Model parameters given in table 2.1.



(a)



(b)

Figure 2-5: (a) The mean and variance of 500 stochastic simulations of Isham's model when a host is treated when its parasite burden exceeds 50 on a discrete set of time periods, and the mean and variance given by a moment closed approximation of the mean and variance without treatment (b) Two examples of the distribution of parasites from stochastic simulations and the approximate distribution using the mean and variance from the moment closure approximations with and without treatment. Model parameters given in table 2.1.

2.3 Attempts to Incorporate the Larval Stage

2.3.1 Simulation Model

When exploring stochastic processes it is often insightful to use simulation models. In preparation for the work that follows, a simulation model was developed using the Gillespie Algorithm [35].

To set up this stochastic model the number of hosts, N, is fixed. The state variables $[L]$ and $[M_i]$ are defined as the size of the free-living larval population, living on a pasture shared by all hosts, and the size of the mature population within the i-th host, for $i \in \{1, \cdots N\}$ at time t. We consider the following processes:

- Larvae are picked up in clumps of size C which has a truncated distribution with the probability generating function $h(z)$. A truncated distribution is used to ensure a host cannot pick up more larvae than exist in the shared larval reservoir.
- The times at which each host will pick up a new clump of parasites are exponentially distributed with rate $\phi[L]/N$, which accounts for the increased likelihood of a host encountering larval parasites as the density of larvae within the area increase, where ϕ is the rate constant.
- Larvae may die as a result of natural causes at rate μ_L per larva.
- Larvae are born at rate β per mature parasite in clumps of size D. This clump size could be a random variable if required but in this model, the size is fixed to simplify the simulations.
- Within a host, parasites may die at rate μ_M per parasite.
- Within a host, parasites may also experience a density-dependent death rate, which acts similarly to the carrying capacity of an environment. Parasites within the i-th host die at an additional rate $D_M[M_i]$ per parasite.

This bears some similarities to the work of Isham and Herbert, extended to consider a truncated distribution on the larval population. Each repetition of the simulation involves a single larval population and N sub-populations of mature parasites. The simulation is carried out using the Gillespie algorithm [35]. The specific implementation used here is as follows:

1. **Initialise system with rate parameters, parameters for clumped distribution, time $t = 0$, and vector containing initial size of larval population and all individual mature sub-populations. Denoted**

$$P = \begin{bmatrix} L(0) \\ M_1(0) \\ M_2(0) \\ \vdots \\ M_N(0) \end{bmatrix}.$$

2. Calculate propensity function as follows

$$a_1 = \beta \sum_{j=2}^{N+1} P_j \qquad \textbf{larval birth}$$

$$a_2 = \mu_L P_1 \qquad \textbf{larval death}$$

$$a_3 = \sum_{j=2}^{N+1} \frac{\phi P_1}{N} \qquad \textbf{any host pick up of larvae}$$

$$a_{3+j-1} = \mu_M P_j + D_M P_j^2 \quad \textbf{for } j \in \{2, \cdots, N+1\} \qquad \textbf{mature parasite death}$$

$$a_0 = \sum_{i=1}^{N+3} a_i \qquad \textbf{total propensity function}$$

3. Generate r_1 and r_2 between 0 and 1 using uniform random number generator.

4. Calculate time at which next event occurs with

$$\tau = \frac{1}{a_0} \ln(\frac{1}{r_1})$$

5. Determine event, labelled k, which occurs using

$$\sum_{i=1}^{k-1} a_i < r_2 a_0 < \sum_{i=1}^{k} a_i.$$

6. If $k = 3$ generate random integer, r_3, between 1 and N from uniform distribution. Generate C from negative binomial distribution truncated at P_1.

7. Update state vector with associated change, changes given by

$$\begin{aligned} k = 1 \qquad & P_1 \to P_1 + 1 \\ k = 2 \qquad & P_1 \to P_1 - 1 \\ k = 3 \quad & P_{r_3+1} \to P_{r_3+1} + C \\ k_{3+j-1} \qquad & P_j \to P_j - 1. \end{aligned}$$

8. Update time with

$$t \to t + \tau$$

9. If $t < T_{final}$ return to step 2. If $t \geq T_{final}$ stop.

If each simulation is indexed by j for W simulations and each of the N hosts within a simulation by i the following may be calculated:

$$\begin{aligned} \overline{L} &= \frac{1}{W}\sum_{j=1}^{W}[L^j] \\ \overline{M} &= \frac{1}{W}\sum_{j=1}^{W}\left[\frac{1}{N}\sum_{i=1}^{N}[M_i^j]\right] \\ V &= \frac{1}{W}\sum_{j=1}^{W}\left[\frac{1}{N}\sum_{i=1}^{N}\left([M_i^j]-\frac{1}{N}\sum_{i=1}^{N}[M_i^j]\right)^2\right]. \end{aligned} \tag{2.12}$$

Using these gives a measure of the mean of the overall distribution of parasites across the linked host-populations, rather than the mean number of parasites in a single host 2-11. The reason that this is interesting is that when compared to the Isham model the distribution of parasites throughout the host population was given by simulating a single population at once and finding the mean and variance of the distribution using

$$\begin{aligned} \overline{M} &= \frac{1}{W}\sum_{j=1}^{W}[M^j] \\ V &= \frac{1}{W}\sum_{j=1}^{W}\left([M^j]-\frac{1}{W}\sum_{j=1}^{W}[M^j]\right)^2, \end{aligned} \tag{2.13}$$

where W is the number of simulations. In effect the mean and variance in the Isham model are taken across the simulations while in the larval model described here the mean and variance are taken within each simulation and then averaged across all simulations.

2.3.2 Model Examples

As the model now has significantly different processes to the Isham model, new parameters must be defined. These are shown in table 2.2. The parameters used for the larval simulation model are quite different from the Isham model. In particular, it should be noted that the time is now measured in months. The model parameters are shown in Table 2.2. They were chosen to keep simulations from becoming too costly; similar results were found across a range of alternative parameter choices. The simulations were run for 50 repetitions. The mean and the variance in the sub-population sizes from each simulation may then be compared with the average mean and variance calculated using (2.12). Figures 2-6,2-7 and 2-8 show comparisons between $\overline{L}$, $\overline{M}$ and V with the larval populations sizes, the mean sub-population sizes and the variance in the sub-population sizes from individual simulations.

Using the assumption that the distribution of parasites throughout the host population is nega-

Parameter/Distribution	Value	Units
$h(z)$	$NB(p,r)$	no units
p	0.8	no units
r	0.5	no units
ϕ	0.03	month^{-1}parasites^{-1}
β	300	month^{-1}
μ_M	1	month^{-1}
μ_L	2	month^{-1}
N	100	hosts
D_M	$\frac{1}{20}$	month^{-1}parasites^{-1}
D	2	parasites

Table 2.2: Parameter values for the simulation model

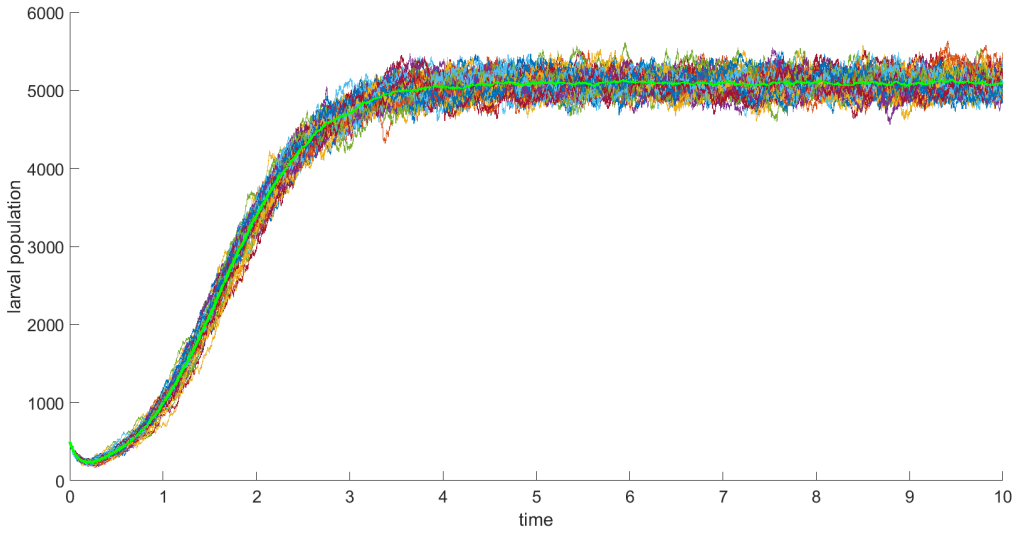


Figure 2-6: Mean of larval populations (green line) compared with the larval populations from the individual simulations

tive binomial, the parameters of this distribution are calculated from the mean and variance in sub-population sizes. Comparisons of the individual simulation p and r values with the p and r values calculated from $\overline{M}$ and V, as defined in (2.12), are given in figures 2-9 and 2-10. Following this, figure 2-11 estimates the distribution of parasites across the host population using the p and r values of a negative binomial distribution calculated from two singular runs of the simulation and by using the average mean and variance. The average distribution gives an estimation that appears qualitatively very similar to those given by estimating from a single simulation of the metapopulation model.

Having established how the stochastic simulation may be used to give an estimation of the distribution of the parasites through the host population figures 2-12 and 2-13 shows what these overall results would look like without the noise of the individual simulation results.

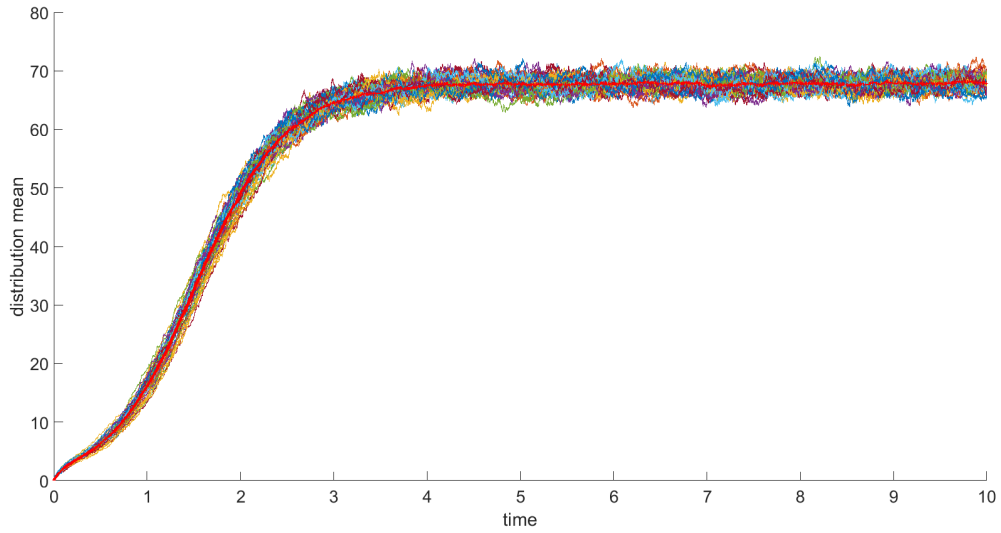


Figure 2-7: Mean of the simulation means (red line) compared with the individual simulation means

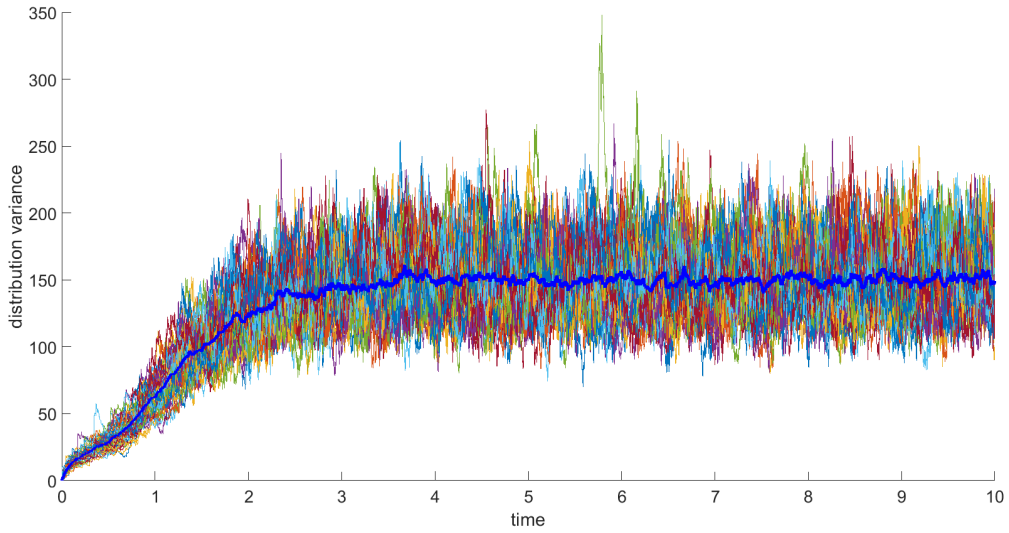


Figure 2-8: Mean of the simulation variances compared with the individual simulation variances, V from (2.12)

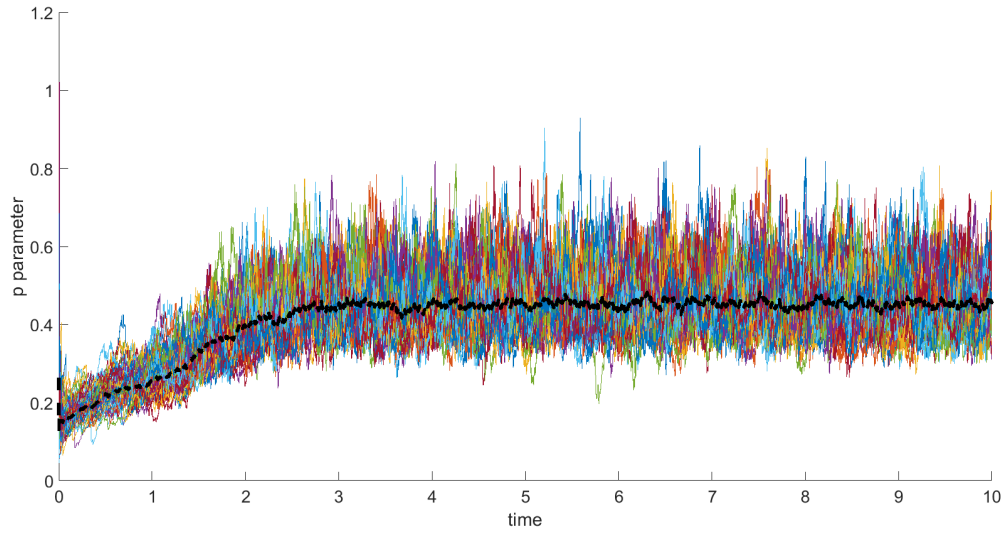


Figure 2-9: "p" parameter of a negative binomial distribution calculated from $\overline{M}$ and V in (2.12)

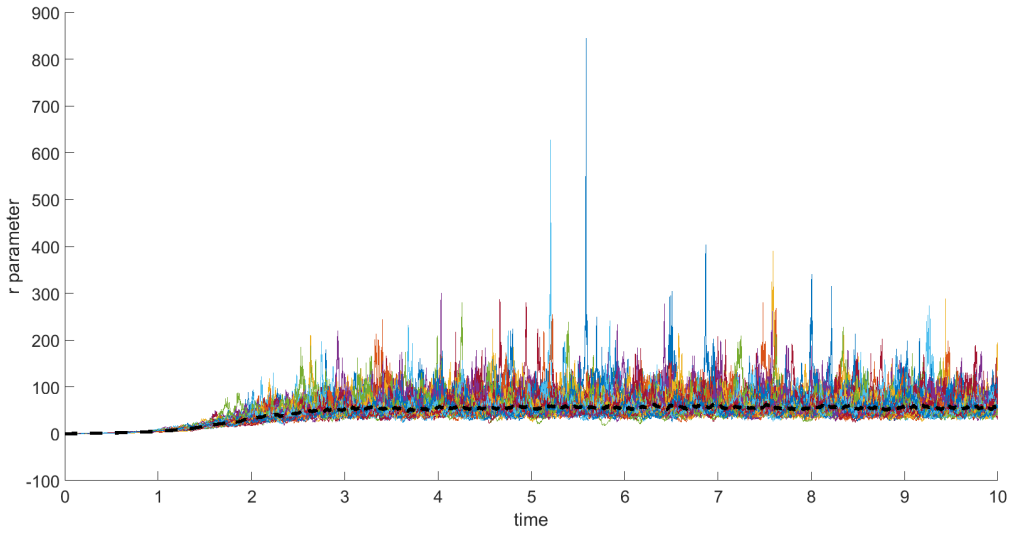


Figure 2-10: "r" parameter of a negative binomial distribution calculated from $\overline{M}$ and V in (2.12)

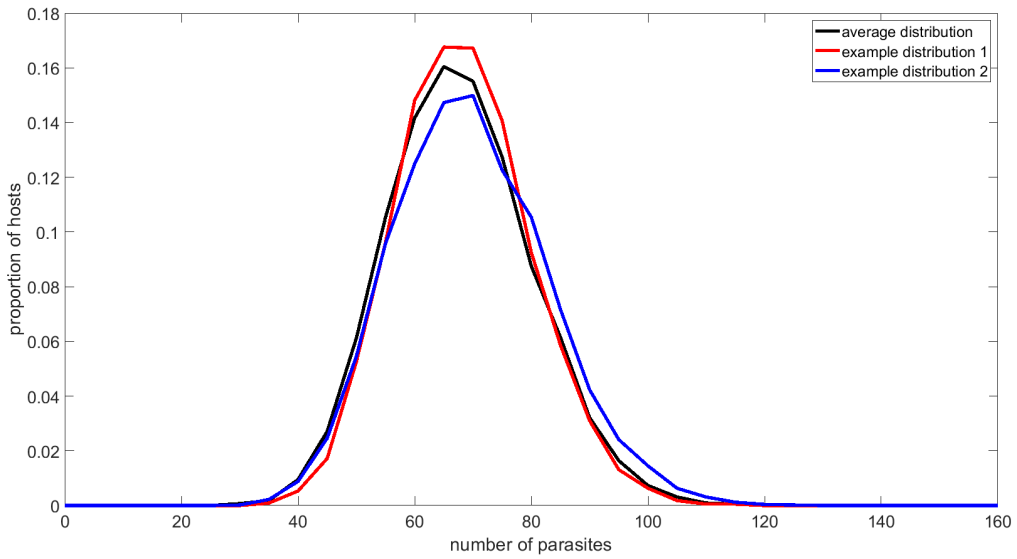
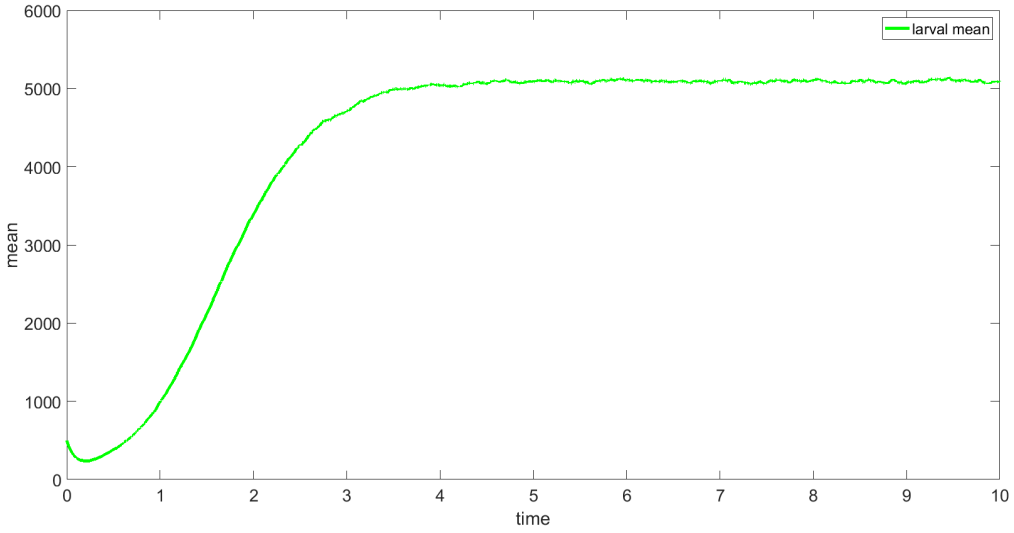
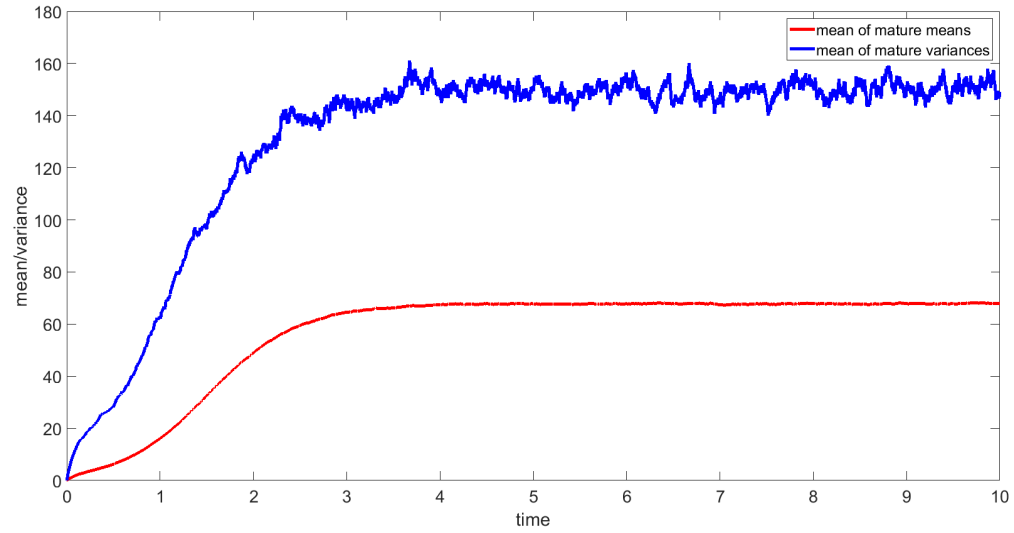


Figure 2-11: Predicted distribution calculated using parameters p and r calculated using (2.9) with mean and variance given by (2.12) ($(\overline{p}, \overline{r}) = (0.46, 57.48)$) compared with examples of the predicted distribution from individual simulations.($(p_1, r_1) = (0.52, 72.69)$ and $(p_2, r_2) = (0.40, 46.84)$)



(a)



(b)

Figure 2-12: Mean across 50 simulations of (a) The larval populations (b) the mean and variances recorded for each simulation taken across the host populations

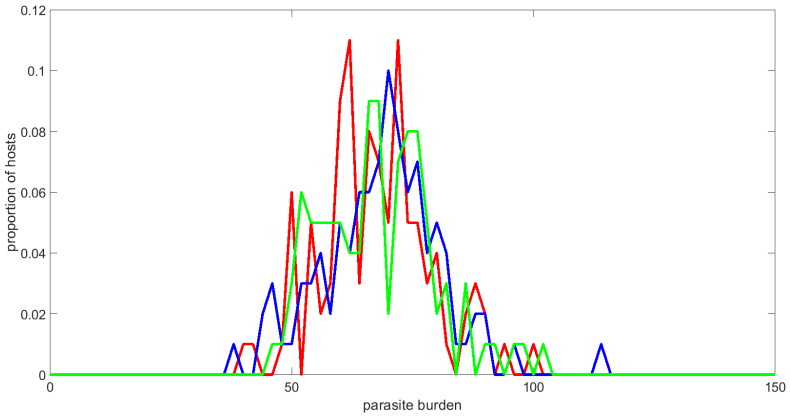


Figure 2-13: Examples of the distribution of parasites through host population at different time points in a single simulation. It should be noted that continuous lines have been used to represent the discrete burdens rather than points to ease the readability of the graph.

2.4 Discussion of Models

Having worked with these models as both fully stochastic simulation models and approximated models this has shown that there are advantages and disadvantages to both approaches. Fully stochastic models allow for treatment to be applied on an individualised basis, as done in section 2.2.2. However, they are computationally costly in comparison to the approximated model. As Isham showed, in certain situations approximated models may perform well and be far less computationally costly, but they may also limiting in how treatment may be applied. For example, it was not possible to treat hosts as soon as their parasite burden exceeded a set level as it was in the stochastic simulations.

In this thesis, the models which are formulated are deterministic approximation models. The work of Isham was very influential in this thesis as the approximated models that are developed in the next chapter are based on a combination of the principles of the stochastic larval model discussed in section 2.3, and the approximating moment closure techniques of the Isham models covered in section 2.1.

Having used Isham's stochastic model to examine individualised treatments we found the key result of this chapter. This is is that when looking at an overdispersed distribution of parasites throughout a host population individualised treatments can have disproportionately larger effects on the variance in the parasite loads than on the mean parasite loads. Going forward with this thesis we aim to try and determine whether these disproportionate effects can alter the optimal control strategies on approximated models.

Chapter 3

Single Species of Parasite Models

3.1 Model Basis

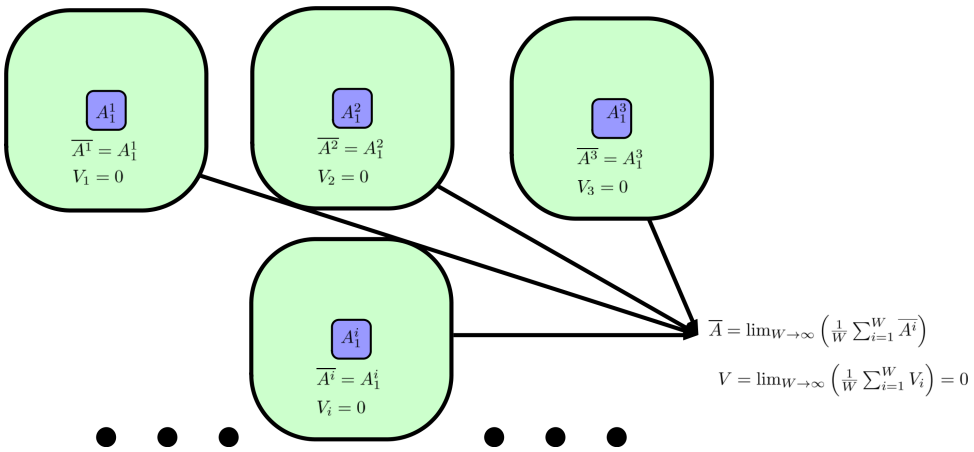
In this chapter, we present a series of model formulations of varying complexity that can be used to describe the population dynamics of a single species of macroparasites. The models described here involve a fixed number of hosts living on a shared pasture with parasites being picked up by hosts, where they then reproduce with the new parasites passing out of the host onto the shared pasture. These larvae may then be picked up by another host and the cycle repeats. Much of the methodology that is detailed within this chapter is based on models devised by Keeling [52] to model metapopulations. Particular effort has been made here to show the progression from the underlying stochastic models to their deterministic counterparts.

3.1.1 Introducing Mean Field Models

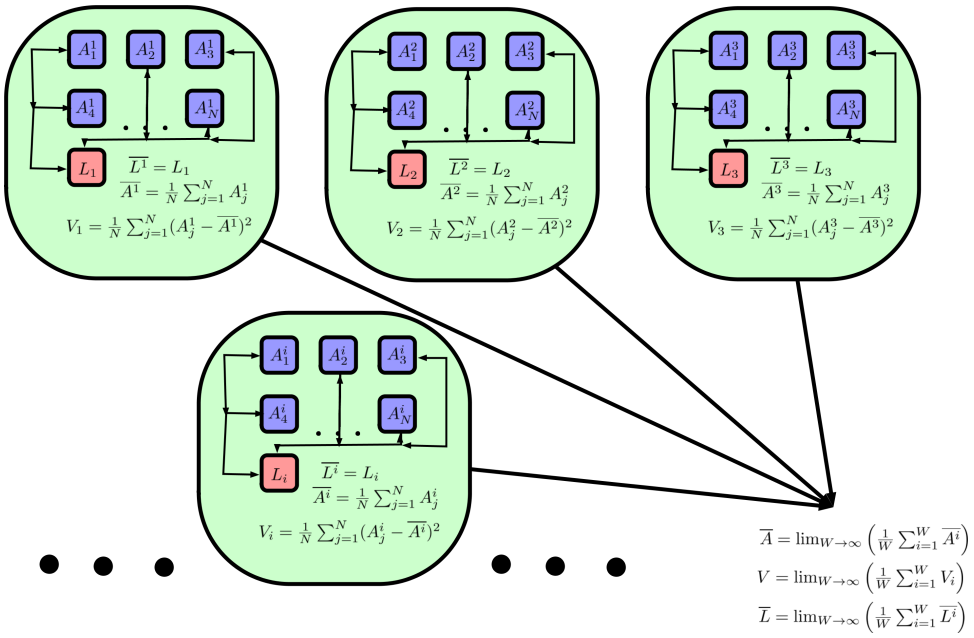
A standard mean-field model for a population assumes that every individual is equally likely to interact with all others in the same way and is subject to the exact same processes as one another. This model can be extracted from a Markov Process by taking the limit as the number of repetitions tends to infinity of the population mean across all simulations [74, 90]. This would give the mean-field approximation of the model.

Figure 3-1 shows a diagrammatic representation of how a standard mean-field approximation differs from the other models derived in this chapter. In the diagrams, each green square represents a single independent simulation of the population or populations being considered, here represented by the purple squares.

In the more relevant linked population models, there are N different populations (corresponding to N hosts) within a single simulation. These populations are linked, through a shared larval population, represented by the red square. For each simulation on the linked population model, there is an



(a)



(b)

Figure 3-1: Diagrams to show the difference in (a) how a standard mean field approximation is calculated and (b) how the mean field approximation is calculated for the linked populations.

associated mean and variance in the size of the parasite population within each host. As figure 3-1 sub-figure b shows this approach creates mean-field models of the metapopulation mean and variance, where the metapopulation consists of the mature parasite populations distributed throughout the N hosts.

3.2 Individual Simulation Model

To formulate the single species models a similar concept to that used in the larval simulation model is used. The mature parasite population within the i-th host during a single simulation, is denoted by $[M_i(t)]$, for $i = 1,\ 2,\ \ldots,\ N$, while the single larval population during a simulation is denoted by $[L]$. With this in mind the variables of interest and the processes to be modelled, within each independent simulation, are as follows:

- Host population size of constant value N.
- Offspring from the parasites in the i-th host are born into the larval population at a rate $f_1([M_i])$ per parasite, where $i \in \{1, 2, \cdots, N\}$.
- Larval parasites are picked up by each host at a rate $f_2([L])$. The probability of an individual larva maturing once in the host is given by ρ; those that do not mature, die.
- Larval parasites that have not yet been picked up die at an average rate of $f_3([L])$ per parasite.
- Mature parasites within the i-th host die naturally at an average rate $f_4([M_i])$ per parasite.
- Mature parasites within the i-th host die due to competition at an average rate $f_5([M_i])$ per parasite.

Here $f_1([M_i]),\ f_2([L]),\ f_3([L]),\ f_4([M_i]),$ and $f_5([M_i])$ all denote the rates of Poisson processes, such that time between events is exponentially distributed [93, 76]. These rates may or may not depend on the overall size of the parasite population in the host at that time. In the model studied here the rates are given by:

$$\begin{aligned} f_1([M_j])) &= \beta \\ f_2([L]) &= \phi \\ f_3([L]) &= \mu_L \\ f_4([M_i]) &= \mu_M \\ f_5([M_i]) &= D_M[M_i]. \end{aligned} \tag{3.1}$$

Here, all parameter values must be greater than zero. The natural parasite death rates, μ_M and μ_L are assumed to be constant, based on the average life span of a mature parasite or a larva in the absence of outside effects. The pick-up rate of larval parasites by hosts, ϕ, assumes that larvae are spread uniformly across the pasture and each host grazes at the same rate and has an equal chance of picking up any larvae. As a result of there being N hosts removing larvae from the pasture the removal rate on the larval population is then ϕN. The death rate due to competition, D_M, acts

somewhat like a carrying capacity of a host, which is justified by the large size of parasites and the limited resources in a host, similar to that of a logistic growth model [74].

Perhaps the most interesting function is the birth rate, $f_1([M_j])$. Parasites often reproduce sexually, and as such require a mate [25]. Given that we do not know the precise number of males and females in a host we seek to find the average birth rate from a host containing $[M_i]$ parasites. In order to do this, we consider a host containing $[M_i]$ parasites. The probability of there being f female parasites and $[M_i] - f$ males in this host is given by

$$\mathbb{P}(f \text{ females }, [M_i] - f \text{ males}) = 2^{-[M_i]} \frac{[M_i]!}{f!([M_i] - f)!}.$$

It is assumed that as long as there is at least one male and at least one female in the population all females will be mated with. As a result, the expected number of parasites producing offspring is given by

$$\sum_{f=1}^{[M_i]-1} \mathbb{P}(f\text{females}, [M_i] - f\text{males}) f.$$

Figure 3-2 shows this expression for increasing values of $[M_i]$. It suggests a a linear relationship between the parasite burden per host and the expected number of mated females may be a reasonable approximation. Therefore we take the birth rate of new parasites as linear with rate constant β.

Having defined the rates and processes we formulate equations for the population size of the larval population, $[L]$, and the mature parasite population in the i-th host, $[M_i]$, at time t. This is done using the conservation equations:

$$[L(t+\delta t)] = [L(t)] + \delta t \sum_{\text{processes on } L} \text{rate of occurence} \times \text{change in } [L] \text{ population caused}$$
$$[M_i(t+\delta t)] = [M_i(t)] + \delta t \sum_{\text{processes on } M_i} \text{rate of occurence} \times \text{change in } [M_i] \text{ population caused.}$$

Using the process described above, this leads to:

$$[L(t+\delta t)] = [L] + \delta t \left(\beta \sum_{j=1}^{N} [M_j](+1) + \mu_L [L](-1) + N\phi[L](-1) \right)$$
$$[M_i(t+\delta t)] = [M_i(t)] + \delta t \left(\rho\phi[L](+1) + \mu_M [M_i](-1) + D_M [M_i]^2(-1) \right).$$

Rearranging and taking the limit as $\delta t \to 0$ results in the rate equations:

$$\frac{d[L]}{dt} = \beta \sum_{j=1}^{N} [M_j] - \mu_L [L] - N\phi[L]$$
$$\frac{d[M_i(t)]}{dt} = \rho\phi[L] - \mu_M [M_i] - D_M [M_i]^2.$$

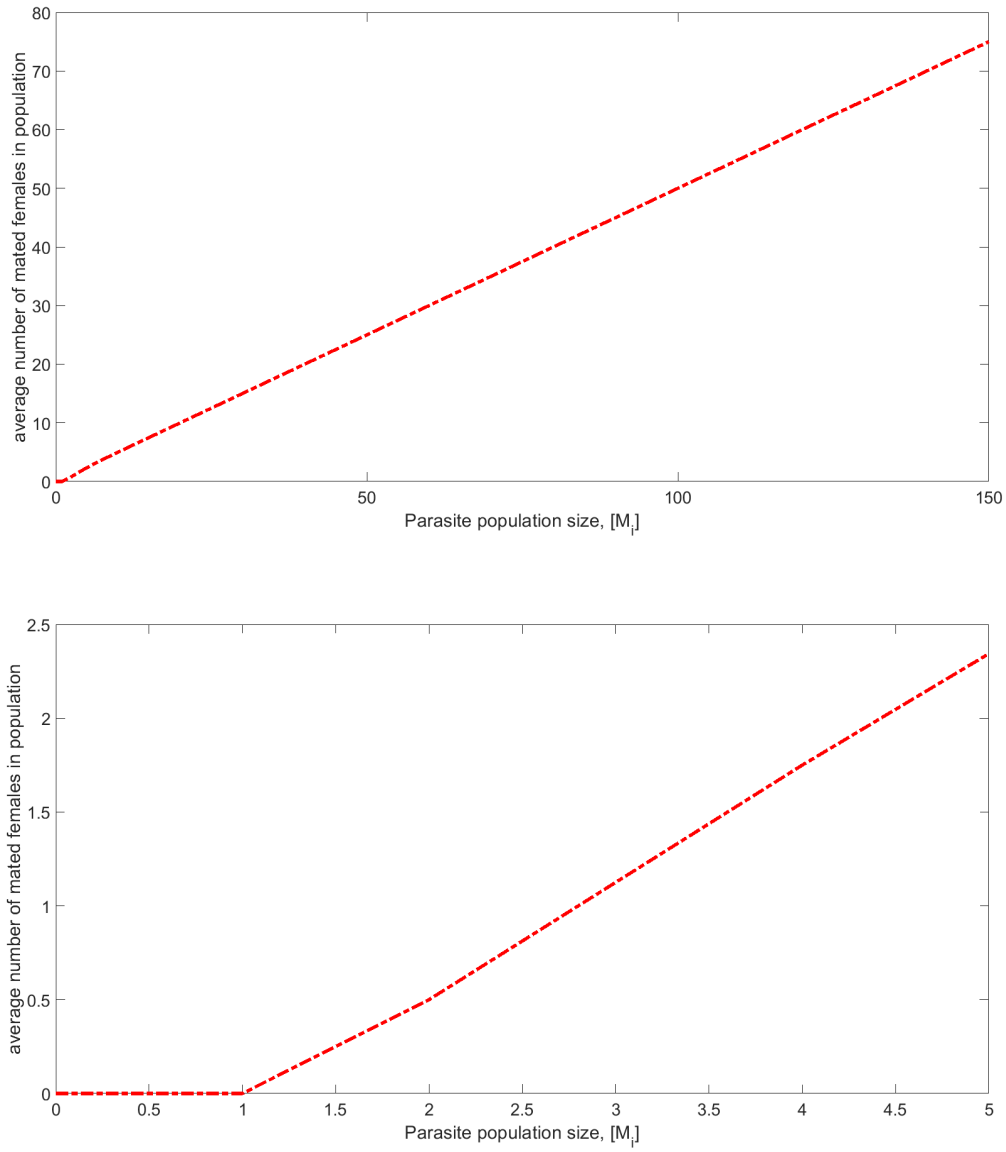


Figure 3-2: (a) Average number of mated females in a host with $[M_i]$ parasites in total (b) Average number of mated females in a host for lower values of $[M_i]$.

Taking the arithmetic mean ($M = \frac{1}{N}\sum_{i=1}^{N}[M_i]$) across the parasite sub-populations in the different hosts gives

$$\begin{aligned}\frac{d[L]}{dt} &= \beta N M - \mu_L[L] - N\phi[L] \\ \frac{dM}{dt} &= \rho\phi[L] - \mu_M M - \frac{D_M}{N}\sum_{i=1}^{N}[M_i]^2.\end{aligned} \tag{3.2}$$

Noting that

$$V = \frac{1}{N}\sum_{i=1}^{N}([M_i] - M)^2 = \frac{1}{N}\sum_{i=1}^{N}[M_i]^2 - (M)^2 \tag{3.3}$$

and substituting into (3.2) gives

$$\begin{aligned}\frac{d[L]}{dt} &= \beta N M - \mu_L[L] - N\phi[L] \\ \frac{dM}{dt} &= \rho\phi[L] - \mu_M M - D_M(M^2 + V).\end{aligned} \tag{3.4}$$

It is important to note that equation (3.4) has been derived for a single simulation, but for ease of notation, no marker was used to indicate this explicitly in the formulation. In the following sections, as we explore how to close this system and incorporate variability due to differences arising from different simulations, we will replace $[L]$ by $[L]^k$, M by M^k and V by V^k for $k = 1, 2, 3...W$ where W is the number of simulations undertaken. As the variance is itself a variable this system is currently unclosed. There are two possible ways that the formulation may proceed:

- Option 1: Use a moment closure approximation which gives the variance in terms of the mean to close the system.
- Option 2: Derive an ODE to describe the time evolution of the variance, which itself will require simplifying assumptions to close the system.

Both these options may be done in a variety of ways by choosing different moment closure approximations. Here the moment closure approximations chosen are selected either for their simplicity or their relevance to the observed distribution of macroparasites in host populations. Option 1 is explored in section 3.3 and option 2 in section 3.4.

3.3 Moment Closure on Variance

If the first option is chosen as a way to close the model there are numerous options for how to approximate the variance by a function of the mean based on different distributions. The simplest of which, which is used here, is to simply set $V^k = 0$, this means that all hosts are assumed to be infected with the same parasite burden. Using this approximation (3.4) becomes

$$\begin{aligned}\frac{d[L]^k}{dt} &= \beta N M^k - \mu_L[L]^k - N\phi[L]^k \\ \frac{dM^k}{dt} &= \rho\phi[L]^k - \mu_M M^k - D_M(M^k)^2.\end{aligned} \tag{3.5}$$

Following the methodology discussed in section 3.1.1 the mean may then be taken across infinite simulations to give $\overline{L} = \lim_{W\to\infty}\left(\frac{1}{W}\sum_{k=1}^{W}[L]^k\right)$ and $\overline{M} = \lim_{W\to\infty}\left(\frac{1}{W}\sum_{k=1}^{W}M^k\right)$. When this is applied to (3.5) this gives

$$\begin{aligned}\frac{d\overline{L}}{dt} &= \beta N\overline{M} - \mu_L\overline{L} - N\phi\overline{L}\\ \frac{d\overline{M}}{dt} &= \rho\phi\overline{L} - \mu_M\overline{M} - D_M(\overline{M}^2 + V_{\overline{M}}),\end{aligned} \tag{3.6}$$

where $V_{\overline{M}} = \lim_{W\to\infty}\left(\frac{1}{W}\sum_{k=1}^{W}(M^k - \overline{M})^2\right)$ is the variance in the means of the simulations. This system is closed by using another moment closure approximation on the variance across simulations, which is chosen as $V_{\overline{M}} = 0$, which is a standard assumption in mean field models. This gives the model as

$$\begin{aligned}\frac{d\overline{L}}{dt} &= \beta N\overline{M} - \mu_L\overline{L} - N\phi\overline{L}\\ \frac{d\overline{M}}{dt} &= \rho\phi\overline{L} - \mu_M\overline{M} - D_M\overline{M}^2.\end{aligned} \tag{3.7}$$

Although we could proceed with this model as is we instead simplify the model by invoking a quasi-equilibrium assumption for the larval population. This is an assumption that has been made by previous authors in the course of building models of macroparasites and is justified by the shorter lifespan of larval parasites than of mature parasites causing processes affecting them to be much quicker [67]. With this assumption, we may set

$$\overline{L} = \frac{\beta N\overline{M}}{\mu_L + N\phi}.$$

Combined with (3.7) gives the simplified model

$$\frac{d\overline{M}}{dt} = \beta_2\overline{M} - \mu_M\beta\overline{M} - D_M\overline{M}^2, \tag{3.8}$$

with $\beta_2 = \frac{\rho\phi\beta N}{\mu_L + N\phi}$. It is on this model that we perform our analysis and impose control.

3.3.1 Analysis of Model

The three properties of greatest interest for these models are:

- The equilibrium values.
- The resilience of the equilibria.
- The reactivity of the steady states.

The equilibrium values are important as they describe the long term state of the system. Additionally, both resilience and reactivity may be dependent on the values that the state variables take at the equilibria. The steady states are determined by setting all equations in the system equal to zero and

solving simultaneously. For a set of ODEs written as $d\mathbf{x}/dt$, this is:

$$\frac{d\mathbf{x}}{dt} = \mathbf{0}.$$

The resilience is very similar to the classic measure of the stability of the equilibria; however, looking at resilience takes the idea of stability and uses it to provide a measure of how quickly the model system would return to equilibrium following a perturbation away from it. For a linear model system given by

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}, \tag{3.9}$$

where A is a matrix, the resilience is given by

$$\text{resilience} = -\Re(\lambda_1(A)), \tag{3.10}$$

where λ_1 is the dominant eigenvalue, the one with the largest real part. Based on (3.10) we see that if the resilience is negative then the steady-state is unstable. While resilience is used as a measure of how quickly a perturbed system returns to an equilibrium, the reactivity, which is sometimes referred to as the initial resilience, is used as a measure of the behaviour directly following a perturbation. It is possible that immediately following a perturbation that a system may show transient growth, where the size of the perturbation initially grows [74, 75]. This is a property of interest in the context of implementing control strategies. For a linear system, a measure of reactivity is given by

$$\text{reactivity} = -\Re(\lambda_1(H(A))), \tag{3.11}$$

where $H(A)$ is the Hermitian matrix given by

$$H(A) = \frac{A + A^T}{2}. \tag{3.12}$$

If the reactivity is positive then the system initially heads towards the equilibrium and transient growth is unlikely. If it is negative, it indicates that the system may exhibit transient growth before settling back to a stable equilibrium level. The greater the magnitude the faster this growth may occur. If $A = A^T$ then the reactivity will be equal to the resilience so for a model of a single state variable, such as the mean-field model that is described in (3.8), transient growth of a perturbed system can not occur around a stable equilibrium. While this may not provide any new information for this model, the value provides a comparison for the next model described in this chapter.

Returning to model system (3.8), we obtain equilibrium values:

$$\overline{M}^* = 0 \quad \overline{M}^* = \frac{\beta_2 - \mu_M}{D_M},\ \beta > \mu_M. \tag{3.13}$$

The condition $\beta_2 > \mu_M$ is imposed to ensure a realistic solution with positive population sizes. Since the properties of resilience and reactivity were defined based on a linear system then an approximated

Equilibria	Resilience	Reactivity	Stability Condition
$\overline{M}^* = 0$	$\mu_M - \beta_2$	$\mu_M - \beta_2$	$\mu_M > \beta_2$
$\overline{M}^* = \frac{\beta_2 - \mu_M}{D_M}$	$\beta_2 - \mu_M$	$\beta_2 - \mu_M$	$\beta_2 > \mu_M$

Table 3.1: Table to show the reactivity and resilience of the trivial and non-trivial equilibrium states of the mean field model

Parameter	Value	Units	Reference
β_2	1.5	days^{-1}	Estimate based on Isham [49]
μ_M	$\frac{1}{5\times 365}$	days^{-1}	[19]
D_M	$\frac{1}{20}$	days^{-1}parasites^{-1}	Estimate

Table 3.2: Parameter values for the basic mean field model (3.8) and basic variance inclusive model (3.24) simulations. Units as required for individual simulation model basis.

system is made based on a linearisation around the steady-state, namely

$$\frac{d(\overline{M}^* + \delta_M)}{dt} = \overbrace{\frac{d\overline{M}^*}{dt}}^{=0} + \frac{d\delta_M}{dt} = \left[\beta_2 - \mu_M - 2D_M\overline{M}^*\right]\delta_M. \tag{3.14}$$
$$\frac{d\delta_M}{dt} = A\delta_M.$$

The eigenvalues of (3.14) are:

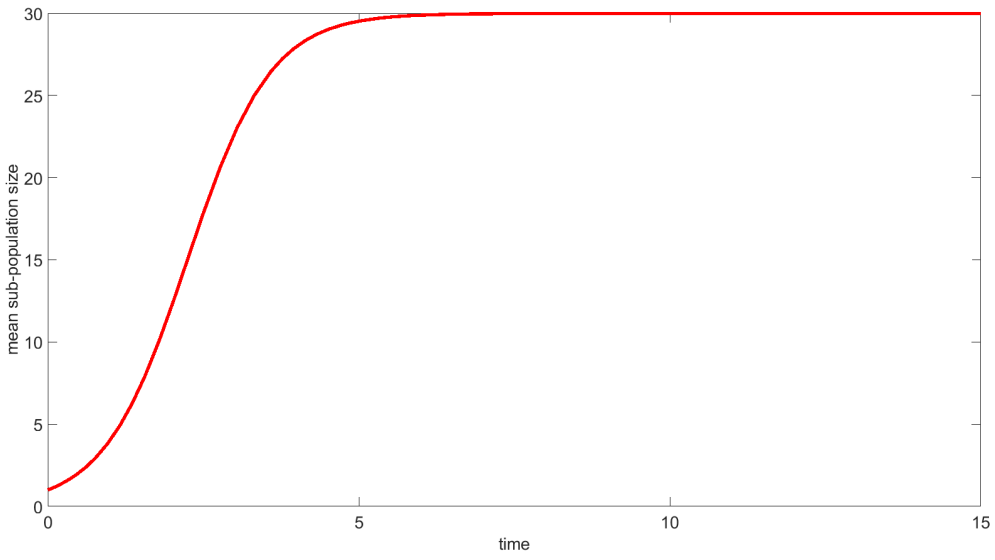
$$\lambda = \begin{cases} \beta_2 - \mu_M & \text{if } \overline{M}^* = 0 \\ \mu_M - \beta_2 & \text{if } \overline{M}^* = \frac{\beta_2 - \mu_M}{D_M},\ \beta_2 > \mu_M \end{cases} \tag{3.15}$$

Under the assumption that $\beta_2 > \mu_M$ then the trivial equilibrium is unstable and the non-trivial equilibrium is stable. Under this assumption, the resilience and the reactivity of the equilibria are as shown in table 3.1. The results in the table lead to the expectation that after a treatment has been applied and the state, $\overline{M}$, perturbed the system will head back to the non-trivial equilibrium without showing any transient growth. The speed at which this occurs will be dependent on the relative values of β_2 and μ_M.

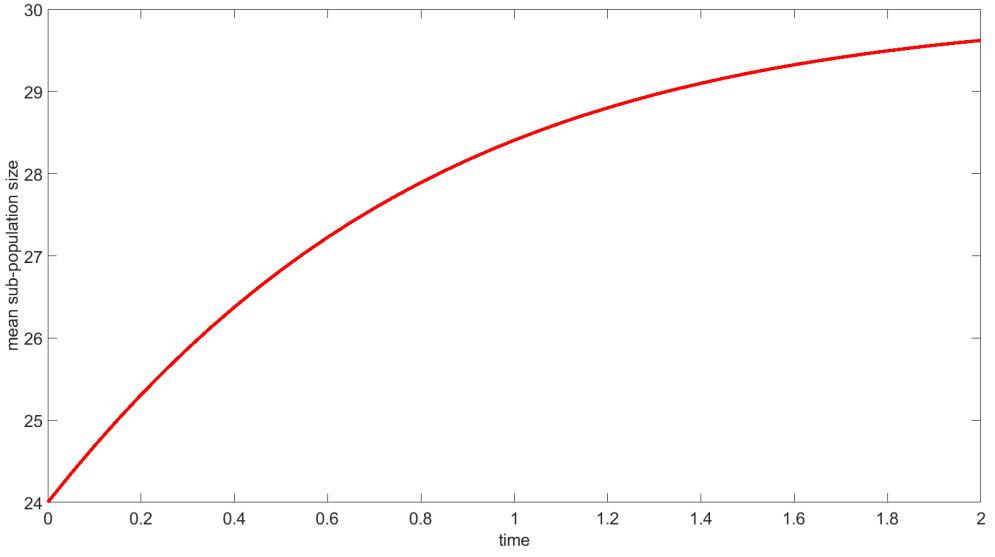
In later models, the complexity of the systems means general formulae for the equilibria, resilience, and reactivity cannot be determined. As a result, conditions under which equilibria are positive, stable, or reactive cannot be found. For comparative purposes, we now focus on a set of parameter values given in table 3.2. These were chosen based on estimations of the lifetime of a parasite within a host, the carrying capacity of a single host, and how many offspring a single parasite would have on average in a day. It must be noted that many of these parameters are not easily determinable for true macroparasitic infections and so are estimated solely to demonstrate model properties. Using these parameters in our model the equilibria, reactivity, and resilience are shown in table 3.3 and the dependent solution moves monotonically to equilibrium as shown in figure 3-3.

Equilibrium	Eigenvalues	Resilience	Reactivity
$\overline{M}^* = 0$	1.4995	-1.4995	-1.4995
$\overline{M}^* = 29.9890$	-1.4995	1.4995	1.4995

Table 3.3: Table to show the reactivity and resilience of the trivial and non-trivial equilibrium states of the mean field model with the established parameters



(a)



(b)

Figure 3-3: (a) Numerical simulation of mean field model heading to equilibrium. (b) Behaviour of the system when state is given small perturbation away from equilibrium. Model parameters given in table 3.2.

3.4 Direct Inclusion of Variance

Returning to the model system from (3.4), and considering a single simulation, we assume now that $V \neq 0$ and not every host is infected to the same extent. Instead, we choose to directly model the variance. Since $V = \frac{1}{N}\sum_{i=1}^{N}([M_i] - M)^2 = \frac{1}{N}\sum_{i=1}^{N}[M_i]^2 - M^2$ [93], then

$$\frac{dV}{dt} = \frac{1}{N}\sum_{i=1}^{N}\frac{d[M_i]^2}{dt} - 2M\frac{dM}{dt}.$$

To determine $\frac{d[M_i]^2}{dt}$, we consider the formula

$$[M_i(t+\delta t)]^2 = [M_i(t)]^2 + \delta t \sum_{\text{processes on } [M_i]} \text{rate of occurence} \times \text{change in } [M_i]^2 \text{ population caused.}$$

When using this formula it is important to note that while the processes and the rate at which they occur is the same for both $[M_i]^2$ and $[M_i]$ the changes in them are not. The changes are shown in table 3.4. Using the rates of change and the change in $[M_i]^2$ shown in table 3.4 we find

$$\frac{d[M_i]^2}{dt} = \rho\phi[L](1 + 2[M_i]) + \mu_M([M_i] - 2[M_i]^2) + D_M([M_i]^2 - 2[M_i]^3). \tag{3.16}$$

To further the model, we take the mean across the N sub-populations, which gives:

$$\begin{aligned}\frac{d(V+M^2)}{dt} = \frac{dV}{dt} + 2M\frac{dM}{dt} &= \rho\phi[L](1+2M^2) + \mu_M(M - 2(V+M^2)) \\ &+ D_M((V+M^2) - 2(T + M^3 + 3MV)),\end{aligned} \tag{3.17}$$

where T is the third order centralised moment and

$$T = \frac{1}{N}\sum_{i=1}^{N}([M_i] - M)^3 = \frac{1}{N}\sum_{i=1}^{N}([M_i]^3) - 3MV - M^3. \tag{3.18}$$

Substituting in the ODE for the mean and rearranging gives the ODE for the variance as:

$$\frac{dV}{dt} = \rho\phi[L] + \mu_M(M - 2(V)) - D_M(2T + 4MV - M^2 - V). \tag{3.19}$$

Due to the presence of the third-order moment, the system is still not closed and once again there are two choices on how to proceed:

Process	Rate	change in $[M_i]$	change in $[M_i]^2$
Production	$\beta\sum_{j=1}^{N}[M_j]$	$+1$	$+2[M_i]+1$
Natural death	$\mu_M[M_i]$	-1	$-2[M_i]+1$
Competition Death	$D_M[M_i]^2$	-1	$-2[M_i]+1$

Table 3.4: Table of rates of process and the change they cause in the size of the parasite populations

1. Use a moment closure approximation,

2. Determine the third-order moment directly.

Modelling the third-order moment would then give a system of equations that would involve the fourth-order moment and the problem would arise again. In this research, the decision was made to use a moment closure to approximate the third-order moment by a function of the mean and the variance. This is also done out of a desire to keep the model simple enough to provide some analytical insights while containing adequate information about the distribution of parasite throughout the N sub-populations.

As discussed previously a phenomenon that has been observed in genuine macroparasitic infections is that the distribution of parasites throughout hosts is far from uniform or even normal. The distributions are typically overdispersed, with a high variance to mean ratio. This results in the majority of hosts being only mildly infected, while a select few are burdened with much higher numbers of parasites. It has been observed that negative binomial distributions may be fitted well to these infections. This combined with the insight gained from Isham's stochastic model is what lead to the decision to use a moment closure approximation based on the negative binomial distribution to close the system [49].

The negative binomial distribution, $NB(r,p)$, is a discrete distribution, originally formulated to model the number of successes that will occur in a sequence of Bernoulli trials before it is stopped. The parameters of the distribution are r, which is the number of failures permitted before the experiment is stopped, and p, which denotes the probability of success in each independent trial [24]. As we have seen it has been used often to model the distribution of parasites throughout a host population [7, 49, 47]. To use the distribution to close our model we use that, for a negative binomially distributed variable, the mean, variance, and skewness are given by:

$$\begin{aligned} M &= \frac{r(1-p)}{p} = \mu_1' \\ V &= \frac{r(1-p)}{p^2} = \mu_2 \\ T &= \frac{r(1-p)(2-p)}{p^3} = \mu_3. \end{aligned} \tag{3.20}$$

Here μ_1' denotes the first order raw moment of the distribution and μ_2 and μ_3 denote the second and third-order centralised moments of the distribution. The raw moments of a distribution, μ_n', are the moments taken about the origin. Defining μ_n' as the n-th order raw moment this is given by:

$$\mu_n' = \mathbb{E}[X^n],$$

where X is the random variable. Similarly, the centralised moments, μ_n, are found by taking the

moments centred around the mean [93], that is:

$$\mu_n = \mathbb{E}[(X - \overline{X})^n], \text{ for } n > 1$$

The second centralised moment denotes V, and the third centralised moment T. The exact relationship between the second and third-order moments is as follows:

$$\begin{aligned} V &= \mu_2 = \mu_2' - (\mu_1')^2 \\ T &= \mu_3 = \mu_3' + 2(\mu_1')^3 - 3\mu_1'\mu_2'. \end{aligned}$$

This may be rewritten in terms of centralised moments to obtain:

$$T = M^2 + (V + M^2)\left[\frac{2(M^2 + V)}{M} - M - 1\right] - M^3 - 3MV. \tag{3.21}$$

The rearrangement of which, and the confirmation of its accuracy, using the formulae from (3.20) is shown below.

$$\begin{aligned} \mu_3' &= \frac{r^2(1-p)^2}{p^2} + \frac{r(1-p)(1+r(1-p))}{p^2}\left[\frac{2r(1-p)(1+r(1-p))}{p} - \frac{r(1-p)}{p} - \frac{p}{p}\right] \\ &= \frac{r^2(1-p)^2}{p^2} + \frac{r(1-p)(1+r(1-p))}{p^2}\left[\frac{r(1-p)}{p} - \frac{2-p}{p}\right] \\ &= \frac{r(1-p)(2-p)}{p^3} + 3\mu_1'\mu_2 - (\mu_1')^3. \end{aligned}$$

When (3.21) is substituted into equation (3.19) and combined with equation (3.4) we obtain the closed system of equations:

$$\begin{aligned} \frac{d[L]}{dt} &= \beta NM - \mu_L[L] - N\phi[L] \\ \frac{dM}{dt} &= \rho\phi[L] - \mu_M M - D_M(V + (M)^2) \\ \frac{dV}{dt} &= \rho\phi[L] + \mu_M(M - 2(V)) \\ &\quad - D_M(\frac{4(V)^2}{M} + 4MV - M^2 - 3V). \end{aligned} \tag{3.22}$$

Recall now that this model has been developed to describe the observations from a single simulation of the host-parasite system with N hosts. To complete the process of using a deterministic system to describe a simulation system, we assume that the variance across sub-population means and

variances are both zero and so we obtain the model system:

$$\begin{aligned}
\frac{d\overline{L}}{dt} &= \beta N\overline{M} - \mu_L\overline{L} - N\phi\overline{L} \\
\frac{d\overline{M}}{dt} &= \rho\phi\overline{L} - \mu_M\overline{M} - D_M(\overline{V} + (\overline{M})^2) \\
\frac{d\overline{V}}{dt} &= \rho\phi\overline{L} + \mu_M(\overline{M} - 2\overline{V}) - D_M(\frac{4\overline{V}^2}{\overline{M}} + 4\overline{MV} - \overline{M}^2 - 3\overline{V}).
\end{aligned} \tag{3.23}$$

Much like the mean-field model, we may assume that the larval population is always at equilibrium and $\frac{d\overline{L}}{dt} = 0$. This gives a simplification of the model, which is only in terms of the mean and variance in the mature parasite burdens of the hosts and is given by:

$$\begin{aligned}
\frac{d\overline{M}}{dt} &= \beta_2\overline{M} - \mu_M\overline{M} - D_M(\overline{V} + (\overline{M})^2) \\
\frac{d\overline{V}}{dt} &= \beta_2\overline{M} + \mu_M(\overline{M} - 2\overline{V}) - D_M(\frac{4\overline{V}^2}{\overline{M}} + 4\overline{MV} - \overline{M}^2 - 3\overline{V}),
\end{aligned} \tag{3.24}$$

with the same formula $\beta_2 = \frac{\rho\phi\beta N}{\mu_L + N\phi}$ as the mean-field model.

3.4.1 Analysis of the Model

Unlike in the mean-field model the variance inclusive model is only valid for $\overline{M} > 0$ due to the inclusion of the $\frac{\overline{V}^2}{\overline{M}}$ term. This does not typically cause a problem as the variance tends to zero at least as fast as the mean. To find the equilibria the ODEs for both the mean and variance are set equal to zero and solved simultaneously,

$$\begin{aligned}
\frac{d\overline{M}}{dt} &= \beta_2\overline{M}^* - \mu_M\overline{M}^* - D_M(\overline{V}^* + (\overline{M}^*)^2) = 0 \\
\frac{d\overline{V}}{dt} &= \beta_2\overline{M}^* + \mu_M(\overline{M}^* - 2\overline{V}^*) - D_M(\frac{4(\overline{V}^*)^2}{\overline{M}^*} + 4\overline{M}^*\overline{V}^* - (\overline{M}^*)^2 - 3\overline{V}^*) = 0.
\end{aligned}$$

Rearranging the ODE for the mean gives the following equation for the variance equilibrium in terms of the mean equilibrium:

$$\overline{V}^* = \left(\frac{\beta_2 - \mu_M}{D_M} - \overline{M}^*\right)\overline{M}^*. \tag{3.25}$$

Setting $Y = \frac{\beta_2 - \mu_M}{D_M} - \overline{M}^*$ this may be written as $\overline{V}^* = Y\overline{M}^*$. Substituting this into the ODE for the variance at equilibrium leads to the following

$$\begin{aligned}
0 &= \beta_2\overline{M}^* + \mu_M(\overline{M}^* - 2\overline{M}^*Y) - D_M(4Y^2\overline{M}^* - 3Y\overline{M}^* - (\overline{M}^*)^2 + 4(\overline{M}^*)^2Y) \\
0 &= \overline{M}^*\left(\beta_2 + \mu_M(1 - 2Y) - D_M(4Y^2 - 3Y - \overline{M}^* + 4\overline{M}^*)\right)
\end{aligned}$$

The simplest solution to this is $\overline{M}^* = 0$, which then leads to $\overline{V}^* = 0$ and gives the trivial equilibrium. The alternative is the solution to the cubic equation in terms of $\overline{M}^*$ given by:

$$(\overline{M}^*)^3(-8D_M) + (\overline{M}^*)^2(4(\beta_2 - \mu_M - D_M)) \\ + (\overline{M}^*)\left(\frac{\beta_2 - \mu_M}{D_M}\left(-2\mu_M - 3D_M - 4\frac{\beta_2 - \mu_M}{D_M}\right)\right) + (\beta_2 + \mu_M) = 0.$$

An alternate option is to consider the non-trivial equilibrium numerically by considering the variance as a function of the mean when the variance equation is set to zero. As the ODE for the variance is quadratic in terms of the variance there will be two possible functions for this. These are:

$$\overline{V} = \overline{M}\frac{\left(2\mu_M + 4D_M\overline{M} - 3D_M\right) \pm \sqrt{\left(2\mu_M + 4D_M\overline{M} - 3D_M\right)^2 + 16D_M(\beta_2 + \mu_M + D_M\overline{M})}}{-8D_M}.$$

Assuming that the $\overline{M} > 0$, the discriminant here is positive, as all parameters are real-valued and positive. The negative root is taken to ensure that the variance is then also positive and admissible. When this equation is satisfied, along with (3.25), then both ODEs will be equal to zero and the system will be at equilibrium. As a result, there will only be one non-trivial equilibrium which describes a realistic state:

$$\left(\overline{M}^*, \overline{M}^*\frac{\left(2\mu_M + 4D_M\overline{M}^* - 3D_M\right) - \sqrt{\left(2\mu_M + 4D_M\overline{M}^* - 3D_M\right)^2 + 16D_M(\beta_2 + \mu_M + D_M\overline{M}^*)}}{-8D_M}\right).$$

When using the negative binomial moment closure the intent was to create a system that would result in a model that could predict the overdispersion that is typically seen in macroparasite population distributions. This is not achieved in the model for any of the parameters sets that were tested. To explore this we first examine what happens when the mean is equal to the variance, an upper bound on the mean for an overdispersed distribution. Setting $\overline{M} = \overline{V} > 0$ in (3.24) leads to

$$\frac{d\overline{M}}{dt} = \beta_2\overline{M} - \mu_M\overline{M} - D_M(\overline{M}^2 + \overline{M}) \\ \frac{d\overline{V}}{dt} = \beta_2\overline{M} - \mu_M\overline{M} - D_M(\overline{M} + 3\overline{M}^2)).$$

The variance equation here may be written as

$$\frac{d\overline{V}}{dt} = \frac{d\overline{M}}{dt} - D_M(2\overline{M}^2).$$

Given that $D_M > 0$ and $\overline{M} > 0$, the derivatives will satisfy $\frac{dV}{dt} < \frac{d\overline{M}}{dt}$ when $V = \overline{M} > 0$. Considering $\frac{dV}{d\overline{M}} = \frac{d\overline{V}}{dt}/\frac{d\overline{M}}{dt} \leq 1$. This means that, if $\frac{d\overline{M}}{dt} > 0$ then $\overline{M}$ will be increasing at a faster rate than V. If $\frac{d\overline{M}}{dt} < 0$ $\overline{M}$ will be decreasing at a slower rate than than $\overline{V}$. In either case these trajectories

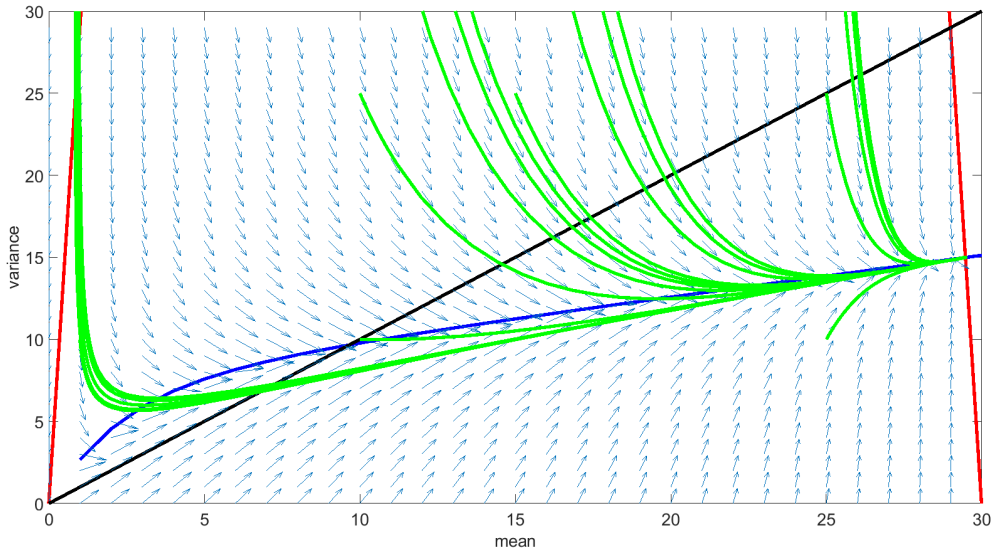


Figure 3-4: Phase plane diagram of the variance inclusive model for parameters shown in table 3.2. Nullclines of the model equations shown in red/blue, trajectories in green and the line $M = V$ in black.

would result in $\overline{M} > \overline{V}$ in the instant following the point at which $\overline{M}$ was equal to V. So for any initial conditions with $\overline{V} < \overline{M}$ an equilibrium reached will also have $V < \overline{M}$ as the variance will be unable to overtake the mean. Next we consider what happens when $\overline{V} > \overline{M}$ by setting $\overline{V} = \overline{M} + \delta$ for $\delta > 0$.

$$\frac{d\overline{M}}{dt} = \beta_2\overline{M} - \mu_M\overline{M} - D_M(\overline{M}^2 + \overline{M} + \delta)$$
$$\frac{d\overline{V}}{dt} = \beta_2\overline{M} - \mu_M\overline{M} - 2\mu_M\delta - D_M(\frac{4\delta^2}{\overline{M}} + 3\overline{M}^2 + 4\overline{M}\delta + \overline{M} + 5\delta).$$
$$\frac{d\overline{V}}{dt} = \beta_2\overline{M} - \mu_M\overline{M} - D_M(\overline{M}^2 + \overline{M} + \delta) - 2\mu_M\delta - D_M(2\overline{M}^2 + 4\delta + 4\overline{M}\delta + \frac{4\delta^2}{\overline{M}}).$$

Once again $2\mu_M\delta + D_M(2\overline{M}^2 + 4\delta + 4\overline{M}\delta + \frac{4\delta^2}{\overline{M}}) > 0$ and $\frac{d\overline{V}}{dt} < \frac{d\overline{M}}{dt}$. This means that either V is decreasing faster than $\overline{M}$ or is increasing slower.

This will mean that at any non-trivial equilibrium it must be the case that $V \leq \overline{M}$ and overdispersion is impossible.

An example of the equilibrium for the variance inclusive model may be seen in figure 3-5. To gain further insight into the model we use that at a non-trivial equilibrium then we will have $0 \leq \overline{V} \leq \overline{M}$. We then use these bounds to try and establish some conditions which may help determine when a non-trivial equilibrium may be stable. If we assume that the variance is at equilibrium then we need

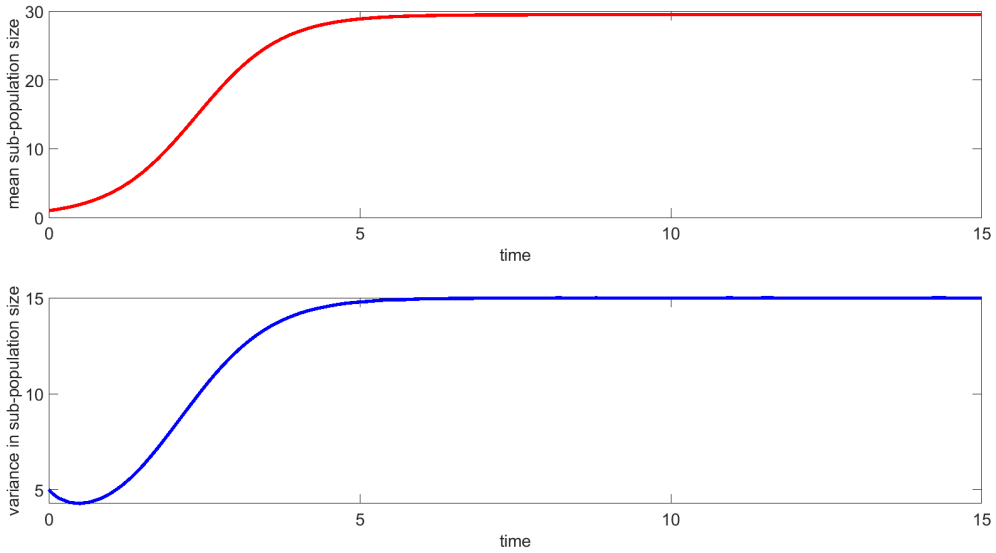


Figure 3-5: Simulation of the variance inclusive model tending to the non-trivial equilibrium for parameters given in table 3.2.

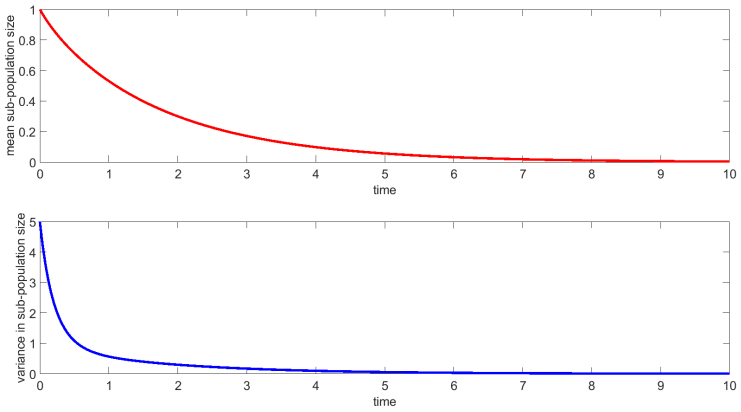
only consider the mean equation. If the value of $\overline{V} \geq 0$ this gives that:

$$\frac{d\overline{M}}{dt} \leq (\beta_2 - \mu_M)\overline{M} - D_M\overline{M}^2.$$

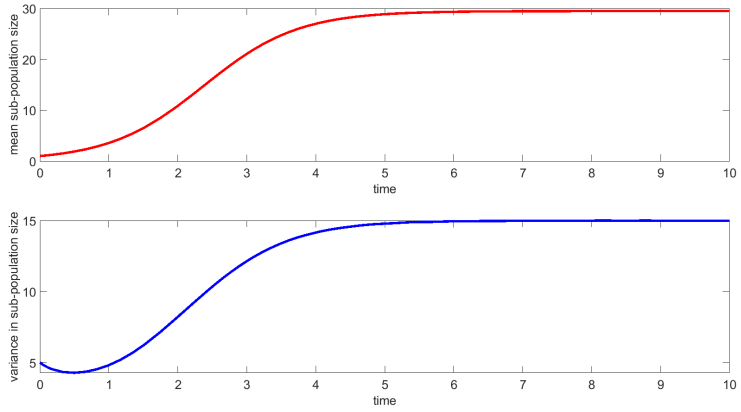
This tells us that the equilibrium value of the mean from the variance inclusive model will be less than or equal to the equilibrium value of the mean from the mean-field model. The mean-field model will tend to the trivial equilibrium if $\beta_2 < \mu_M$ and so the same should be true for the variance inclusive model. If we consider the other bound on $\overline{V}^*$, that is set $V \leq \overline{M}$ then we find:

$$\frac{d\overline{M}}{dt} \geq (\beta_2 - \mu_M - D_M)\overline{M} - D_M\overline{M}^2.$$

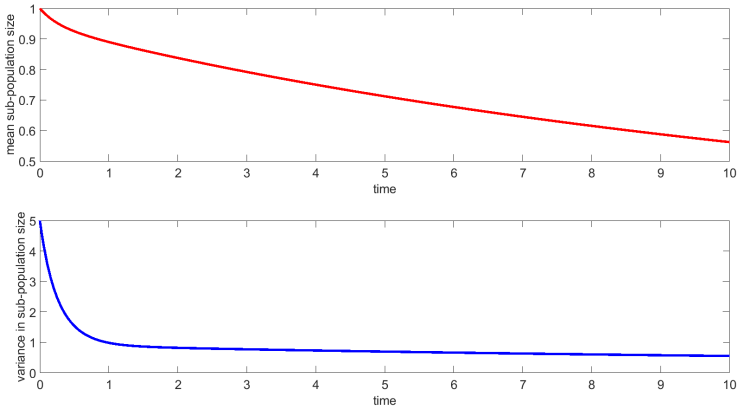
If we consider when the left-hand side of this equation would be equal to zero we find that $\overline{M} = \frac{\beta_2 - \mu_M - D_M}{D_M}$, which will have a positive value if $\beta_2 > \mu_M + D_M$. This suggests that under this same condition the variance inclusive model will have a non-trivial equilibrium state with $\overline{M}^* \geq \frac{\beta_2 - \mu_M - D_M}{D_M}$ and $0 \leq \overline{V}^* \leq \overline{M}^*$ which has the potential to be stable. When $\mu_M < \beta_2 < \mu_M + D_M$ it is unclear what the outcome will be. We consider three parameter sets for these conditions to numerically determine the precise equilibria and their stability. The parameter sets are given in table 3.5. The simulation of the model may be seen for each of the three parameter sets in figure 3-6. When $\mu_M > \beta_2$ the parasites die out and when $\beta_2 > \mu_M + D_M$ the non-trivial equilibrium is reached. When $\mu_M < \beta_2 < \mu_M + D_M$ then the outcome appeared to be that the parasites persisted at a very low level the variance was extremely close to the mean. To understand it better we consider the eigenvalues of the approximate equilibria from the numerical determination under each of these conditions. These may be seen in table 3.6. What these show is that, when $\beta_2 < \mu_M$, as predicted the trivial state is stable. In the case that $\beta_2 > \mu_M + D_M$ we see that the trivial state is unstable as



(a)



(b)



(c)

Figure 3-6: Simulation of the variance inclusive model with parameter set such that (a) $\mu_M > \beta$ (b) $\mu_M + D_M < \beta$ (c) $\mu_M < \beta < \mu_M + D_M$

Parameter set	Parameter	Value
$\beta_2 > \mu_M + D_M$	μ_M	$\frac{1}{5\times 365}$
$\beta_2 < \mu_M$	μ_M	2
$\mu_M < \beta_2 < \mu_M + D_M$	μ_M	$1.5 - \frac{1}{30}$
All	β_2	1.5
All	D_M	$\frac{1}{20}$

Table 3.5: Parameter sets for the three conditions which may affect the stability of the non-trivial equilibrium for the variance inclusive model.

.

Parameter set	Equilibrium	Eigenvalues of A	Eigenvalues of $H(A)$
$\mu_M > \beta_2$	$(0, 0)$	$-3.7969,\ -0.5531$	$-4.5794,\ 0.2294$
$\beta_2 > \mu_M + D_M$	$(29.4804,\ 14.9895)$	$-1.4653,\ -5.9338$	$-6.0648,\ -1.3344$
$\beta_2 > \mu_M + D_M$	$(0, 0)$	$1.4414,\ 0.2070$	$1.8151,\ -0.1668$
$\mu_M < \beta_2 < \mu_M + D_M$	$(0, 0)$	$-2.7296,\ -0.0204$	$-3.4023,\ 0.6523$

Table 3.6: Eigenvalues of the equilibria for each of the three parameter sets given in table 3.5.

both eigenvalues are real and positive but the approximate equilibrium from the numerical simulations is shown to be stable. Although the results of simulating the case where $\mu_M < \beta_2 < \mu_M + D_M$ show a very low-level persistence of the parasites, it is clear that the final equilibrium has not been reached. The eigenvalues however suggest that the parasites will eventually die out, although given the values of these eigenvalues it will take a significant amount of time (resilience = 0.0204, corresponding to a time of 49 time units to reach equilibrium). However, even when simulations were run for longer than this the trivial equilibrium was not reached.

3.5 Conclusions

In this chapter, we have formulated two models based on the same stochastic basis for the distribution of parasites throughout a host population of fixed size. The first model assumes all hosts have the same parasite burden with no variance. The second assumes an underlying negative binomial distribution for the parasite burdens, which is used to close the model using a moment closure approximation for the third-order moment.

In the analysis of our first model, we were able to show that under a single condition on the size of the birth rate and parasite death rate, $\beta_2 > \mu_M$, that a positive non-trivial equilibrium would exist and be stable. This provided a condition that we used to ensure that the model simulations do not tend to the trivial equilibrium and make treatment unnecessary in the following chapter. In examining the reactivity of the equilibria we found that the stable equilibria would be non-reactive, and transient growth following perturbations would be unexpected. This result provides a point of comparison with later models.

The second model we formulated provided multiple important results. The first of which was that,

despite using a moment closure approximation based on the relationship between the third-order moment, and the first and second order moments of a negative binomial distribution, this was not enough to lead to an aggregated distribution in the ODE model. By comparison with the Isham model, we theorise that to create a model that would exhibit overdispersion, then the variance would need to be included elsewhere in the model system, not simply arising from a density-dependent death term which by its very nature is stabilising and acts to reduce variability. The second important result was that when compared with the first model we needed to consider three conditions that would affect the existence and stability of a positive equilibrium. If the birth rate was less than the mature parasite death rate only the trivial equilibria was viable and stable; if the birth rate was greater than the mature death rate and the density-dependent death rate a positive equilibrium was shown to exist and have stability. The third potential case where $\mu_M < \beta_2 < \mu_M + D_M$ we found that no equilibrium was reached in simulations. These conditions were then used to set example parameters for the control problems seen in the following chapter.

Chapter 4

Application of Control

This chapter begins to consider the issue of control on a macroparasitic infection across multiple hosts. This begins with an examination of the problem formulated as a standard optimal control problem with a continuous control applied across the host population. This is intended to allow for a more direct comparison with how a problem such as this may be formulated and optimised on a typical mean-field model for microparasitic diseases. It is also a formulation that may have more general applicability for models formulated in a similar way for linked populations where control intensity can be effectively changed. Following this, the problem of control is looked at as a binary control problem, where control may be turned off or on at a fixed intensity [88]. This is one of the most crucial aspects of the control problem on our macroparasitic models where treatment dosages are typically set and the decision is a choice of if and when to treat [4, 51]. These ideas have been discussed somewhat in relation to the Isham models in Chapter 2; in this chapter, we focus on how models from Chapter 3 may be used for comparing treatment strategies and how the optimal control may be determined. For comparative purposes we also return to the Isham model, in a more formal framework, to look in greater depth at how an optimal control may differ on a model in which over dispersal is achieved.

4.1 Basic Optimal Control

Standard optimal control problems take a dynamical system and apply control to it. The aim of the control is typically to minimise or maximise some associated cost function, $J(u, x, t)$, where $x(t)$ and $u(t)$ denote the state of the dynamical system and the control, respectively, at time t. The control function, $J(u, x, t)$, may also be used to impose adherence to an end condition, $L(x, u, T_f)$ [58]. The methodology by which optimal controls may be determined uses Pontryagin's 'maximum principle[81, 15]. The following section discusses this principle for a general dynamical system.

4.1.1 Pontryagin's Principle

We define a general dynamic system of the form

$$\frac{dx_i}{dt} = f_i(\underline{x}, \underline{u}, t) \qquad \text{for } i = 1, \ldots, n,$$

where $\underline{x} = (x_1, x_2, \ldots, x_n)$ is the state of the system, $\underline{u} \in \mathcal{U}$ is a control input which belongs in the set of admissible controls for the problem, and the system has initial condition $\underline{x}(0) = \underline{x}_0$. The aim of control is to minimise a cost function,

$$J(\underline{x}, \underline{u}, t) = \int_0^{T_f} f_0(\underline{x}, \underline{u})dt + L(\underline{x}(T_f)). \tag{4.1}$$

This function comprises of a cost based on the final state, $L(\underline{x}(T_f))$, and a cost based on amount of control effort applied and the system state at every point in the control interval, $\int_0^{T_f} f_0(\underline{x}, \underline{u})dt$.

Pontryagin's principle states that an optimal control, u^*, and the corresponding optimal state trajectory, $\underline{x}^*$, and co-states (Lagrange multipliers associated with the state equations), $\underline{\lambda}^*$, must minimise the Hamiltonian associated with the system in comparison to all other admissible controls [81]. That is

$$\mathcal{H}(\underline{u}^*, \underline{x}^*, \underline{\lambda}^*) \leq \mathcal{H}(\underline{u}, \underline{x}, \underline{\lambda}).$$

To make use of this principle we must first define the co-state equations. For the general system given above, these are given by:

$$\begin{aligned} \lambda_0 &= 1 \\ \frac{d\lambda_i}{dt} &= -\sum_{j=0}^{n} \frac{\partial f_j}{\partial x_i}\lambda_j, \text{ for } i = 1, \ldots, n. \end{aligned}$$

For a particular trajectory these may be found using the final state condition that $\lambda_i(T_f) = \frac{\partial L(\underline{x}(T_f))}{\partial x_i}$, determined from the cost function above. With these defined then the Hamiltonian may be defined as

$$\mathcal{H}(\underline{u}, \underline{x}, \underline{\lambda}) = \sum_{i=0}^{n} f_i(\underline{x}, \underline{u}, t)\lambda_i. \tag{4.2}$$

As the admissible controls are assumed to be piecewise continuous the Hamiltonian may be differentiated with respect to $\underline{u}$ to give the optimality condition that

$$\frac{\partial \mathcal{H}(\underline{u}^*, \underline{x}^*, \underline{\lambda}^*)}{\partial \underline{u}} = 0,$$

which must be satisfied when a control and its associated state trajectories and co-states minimises the Hamiltonian. By rearranging this a function for $\underline{u}^*$ in terms of $\underline{x}^*$ and $\underline{\lambda}^*$ is determined. When all admissible controls are bounded this condition may not be satisfied. When this is the case the optimal control, within the set of admissible controls, may still be determined by considering the

Parameter	Value	Units	Definition
β_2	1.5	[days]$^{-1}$	Birth/pick up rate of new parasites
μ_M	$\frac{1}{5\times 365}$	days^{-1}	Natural death rate
D_M	$\frac{1}{20}$	days^{-1}parasites^{-1}	Death rate due to competition
u_{max}	1	days^{-1}	Maximal increased death rate due to treatment

Table 4.1: Model parameters for the mean-field model control examples

value of the partial derivative and its relation to the aim of the control. For a minimisation problem on a set of bounded controls, the optimal control is given by

$$u_j = \begin{cases} u_{max} & \text{if } \frac{\partial \mathcal{H}}{\partial u_j} < 0 \\ u_{min} & \text{if } \frac{\partial \mathcal{H}}{\partial u_j} > 0 \\ u_j^* & \text{if } \frac{\partial \mathcal{H}}{\partial u_j} = 0, \end{cases} \text{ , for } j = 1, \cdots, m.$$

Here u_j denotes one of the control inputs. For a maximisation problem, the signs are reversed [58].

Except for the simplest linear cases, explicit determination of the optimal control is not possible and so numerical methods must be employed [58]. For the optimal control problems here, BOCOP software was used to solve for the optimum control. The BOCOP software uses an algorithm that converts the optimal control problem into a finite-dimensional optimisation problem by discretising the state and control over time. This allows the problem to be formulated as a non-linear programming problem which is solved via a primal-dual interior-point algorithm. This algorithm involves the computation of approximate solutions to a sequence of barrier problems that have a decreasing sequence of barrier parameters [32]. The solution of a barrier problem is then used to continue the solution for the next problem in the sequence. When an error estimate is below a set tolerance then the algorithm stops. This algorithm is available as an open-source solver IPOPT, which the BOCOP software utilises to solve the barrier problems after it has converted the specified problem into the correct form [95, 16]. The definition files for use with the BOCOP software for each of the examples given in this, and later chapters, are linked in the appendices (see section C).

4.1.2 Control Example Parameters

To maintain continuity throughout the examples that we show here we will use the same model parameters for the different types of control examples. These parameters for the mean-field model are given in table 4.1, the variance inclusive model in table 4.2 and the Isham model in table 4.3. The maximal control parameter, u_{max}, is chosen to allow significant reductions in the macroparasite burden. It would not, however, result in the eradication of the parasites to reflect realistic scenarios in which treatment reduces but does not eradicate burdens across a host population.

The parameters for the mean-field and variance inclusive model are the same as in the previous chapter, with the addition of the maximal treatment parameter u_{max}. The Isham model parameters

Parameter	Value	Units	Definition
β_2	1.5	[days]$^{-1}$	Birth/pick up rate of new parasites
μ_M	$\frac{1}{5\times 365}$	days^{-1}	Natural death rate
D_M	$\frac{1}{20}$	days^{-1}parasites^{-1}	Death rate due to competition
u_{max}	1	days^{-1}	Maximal increased death rate due to treatment

Table 4.2: Model parameters for the variance inclusive model control examples.

Parameter	Value	Units	Definitions
ϕ	$\frac{52}{12}$	month^{-1}	Average rate clumps are picked up
μ_H	0	month^{-1}	Rate of host death from non-parasite causes
μ_M	$\frac{1}{12}$	month^{-1}	Death rate of each parasite
α	0.002	month^{-1}parasites^{-1}	Increase in host death rate due to a single parasite
p	0.5	no dimensions	Parameter of clump distribution
r	0.5	no dimensions	Parameter of clump distribution
u_{max}	1	month^{-1}	Maximal increase in parasite death rate from control

Table 4.3: Model parameters for the Isham model used in all treatment examples

are altered such that it has an untreated equilibrium mean close to those of the other two models.

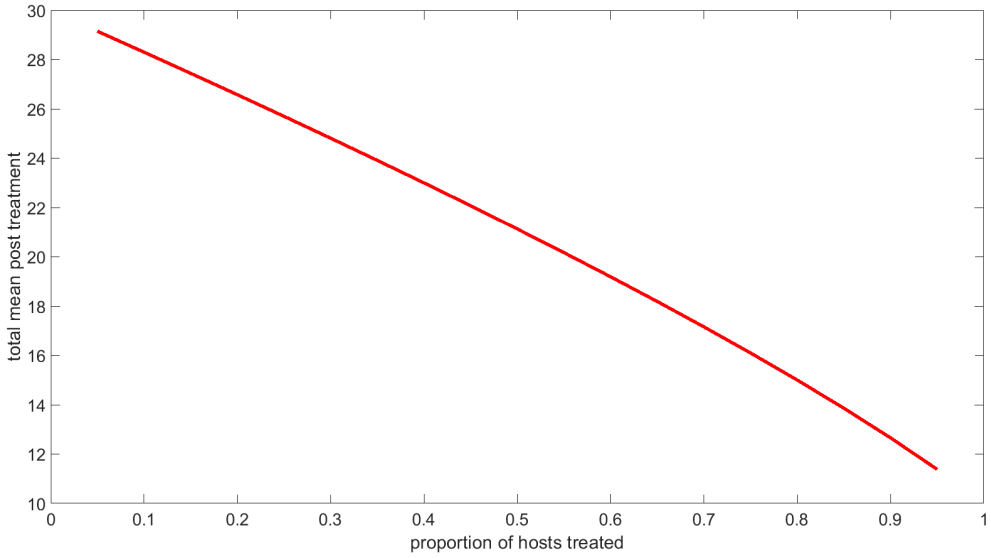
4.1.3 Application to the Mean-Field Model

This section first sets up an optimal control problem on the mean-field model from (3.8) and then examines how the control behaves in this case. This is then used for comparative discussion with the variance inclusive model.

We choose to examine a simple control where treatment is applied to a host causing an increase in the linear death rate that the parasites suffer. This increase in the linear death rate may be caused by numerous effects of treatment, such as through agitation of neuro-receptors of parasites causing paralysis [34]. Other potential effects of treatments that we do not consider, may cause effects on different parameters, such as inhibition of parasite absorption which could lower the carrying capacity of a host and resultantly increasing D_M [64].

Control Applied to All Hosts

If we assume, the control is applied to all hosts at the same intensity and causes an increase in the linear death rate μ_M, then the only change in the individual simulation model is that the linear death rate μ_M is replaced by $\mu_M + u(t)$. As a result, the model is essentially the same, as (3.8) for the mean-field model or (3.24), just with this new rate parameter in place.



(a)

Figure 4-1: (a) Value of $M^* = M_U^* + M_T^*$ when treatment is applied to different proportions of the host population for a sustained period of 100 units 4.1.

Control Applied Selectively

If we examine the possibility that our control variable is representing a proportion of hosts that are treated the individual simulation model changes far more. Rather than a single population of hosts, we would have two populations; a treated population, and an untreated population. To prevent confusion we change our control variable to p for this example with p being the proportion of hosts undergoing treatment. We first assume that p is fixed so that our treated population size is pN and our untreated population size is $(1-p)N$. The two host populations share a living space so that parasites born within one host may go into any other host. By assuming that there is no variance in the parasite population in the two groups of hosts (treated and untreated), using (equation (3.8)), we obtain the coupled system

$$\begin{aligned} \frac{dM_U}{dt} &= \beta_2 N\left(pM_T + (1-p)M_U\right) - \mu_M M_U - D_M M_U^2 \\ \frac{dM_T}{dt} &= \beta_2 N\left(pM_T + (1-p)M_U\right) - (\mu_M + \mu_T)M_T - D_M M_T^2 \end{aligned} \tag{4.3}$$

where $M_U(t)$ and $M_T(t)$ denote the hosts that are untreated and treated at time t respectively. The equation for the overall mean is simply equal to the proportional addition of the two separate means, that is

$$\frac{dM}{dt} == p\frac{dM_T}{dt} + (1-p)\frac{dM_U}{dt}.$$

Additionally, changing the proportion of hosts being treated, in this way, will introduce discontinuity to the model, as a host moves between categories, at the time the proportion changes. This

Parameter	Value
α_1	$\frac{1}{11.74^2}$
M_T	9.5

Table 4.4: Parameters of the cost function (4.5) for the mean-field model

discontinuity makes this selective treatment a poor candidate for optimisation as any change in the treatment variable would require the instantaneous changes to the treated and untreated means to be calculated. Moving a host from the treated group to the untreated group, or vice versa, would also introduce variance as the burden of the host moving categories will not be equal to the mean burden of the new category. This would violate the model assumption that the variance is zero and the model would be unusable.

Formulating the Control Problem

When $p = 1$, every host is subjected to treatment at the same intensity u, as discussed above. If this treatment intensity is now used as the control variable, which may be changed continuously within $0 \leq u(t) \leq u_{max}$, the dynamics of the model under treatment become:

$$\frac{dM}{dt} = \beta_2 M - (\mu_M + u(t))M - D_M M^2, \qquad \text{for } 0 \leq u(t) \leq u_{max}. \tag{4.4}$$

with associated cost function

$$J(u) = \int_0^{T_f} u(t)^2 + \alpha_1 (M - M_T)^2 dt, \tag{4.5}$$

where α_1 is a positive constant, $M_T \geq 0$ is the target and T_f denotes the time-period over which the control is applied. The reason for this is that we wish our control to balance the cost of control with the cost associated with the state variable $\overline{M}$. We follow standard approaches to optimal control by using quadratic terms in the objective function. However, it is also appropriate since it accurately reflects our control goal to minimise the amount of treatment (for cost and well-being reasons) whilst also reducing parasite burden to a low target level.

The cost of control is normalised so that the only cost parameter is α_1. We choose $\alpha_1 = \frac{1}{(M_x - M_T)^2}$ such that the relative cost of treatment and the cost of parasite burden above the target M_T are equal at a given parasite burden, M_x. Similar cost functions are used throughout this thesis. For controls with this cost function on the mean-field model we use the cost function parameters given in table 4.4. These parameters are chosen such that M_T is approximately equal to the equilibrium value of the mean when full treatment is applied. The value of α_1 is chosen such that the value of M_x would be halfway between the fully treated and fully untreated equilibrium values of the mean. In the absence of any data, decisions about the size of α_1 have to be made; our approach has been to make a choice that can be justified in the context of our theoretical study. Whilst quantitative differences would arise due to different choices, the nature of this control problem means that qual-

itatively similar results would be found with other, less clearly defined choices of α_1.

With the dynamic system and the cost function defined we obtain the Hamiltonian:

$$H(u, \overline{M}) = u^2 + \alpha_1(\overline{M} - M_T)^2 + \lambda_1 \left(\beta_2 \overline{M} - (\mu_M + u)\overline{M} - D_M \overline{M}^2\right), \tag{4.6}$$

the governing equation for λ_1:

$$\frac{d\lambda_1}{dt} = -2\alpha_1(\overline{M} - M_T) - \lambda_1(\beta_2 - \mu_M - u - 2D_M\overline{M}), \qquad \lambda_1(T_f) = 0.$$

The condition of optimality is:

$$\frac{\partial \mathcal{H}(u^*, \overline{M}^*)}{\partial u} = 2u^* - \lambda_1 \overline{M}^* = 0.$$

As we have imposed upper and lower bounds on the admissible controls, we need further conditions on this to give an optimal control that meets these conditions [58]. In this case, the optimality condition is given by

$$u^* = \begin{cases} \frac{\lambda_1 \overline{M}^*}{2} & \text{if } \frac{\partial \mathcal{H}}{\partial u} = 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0 \\ 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0. \end{cases}$$

To solve the optimal control problem we used the BOCOP software, the output of which is shown in figure 4-2 [16]. This figure shows that the maximal intensity control is applied until the state reaches approximately $M_T + M_x = 9.5 + 11.74$ as intended. Past this point, it quickly settles to a lower intensity. At the end of the interval, the control is stopped as the cost function does not penalise a higher end condition and the optimisation does not need to account for what happens after. Despite the control parameter α_1 being set such that M_x is halfway between M_T and the untreated equilibrium value of the model, the control intensity is substantially greater than a half of the maximum potential control.

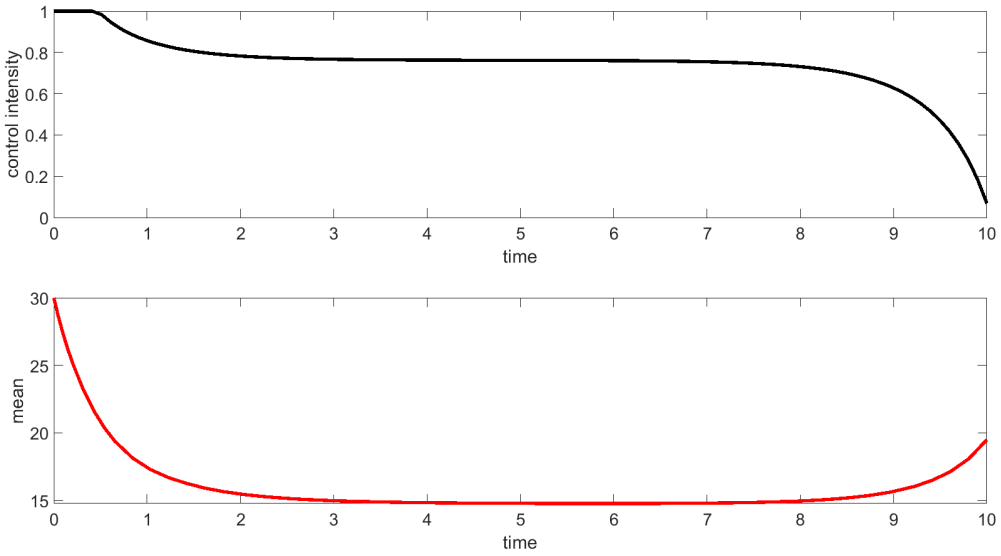


Figure 4-2: Continuous optimal control on the mean-field model (3.8) showing the control and the mean population size as a result of the control. Parameter values in table 4.1 and cost parameters in table 4.4, mean-field model.

4.1.4 Application to the Variance Inclusive Model

Assuming that there is variation within the host's parasite burdens then (4.3) becomes:

$$\begin{aligned}\frac{d\overline{M}_U}{dt} &= \beta_2\left(p\overline{M}_T + (1-p)\overline{M}_U\right) - \mu_M\overline{M}_U - D_M(\overline{M}_U^2 + V_U)\\ \frac{d\overline{M}_T}{dt} &= \beta_2\left(p\overline{M}_T + (1-p)\overline{M}_U\right) - (\mu_M+\mu_T)\overline{M}_T - D_M(\overline{M}_T^2 + V_T).\end{aligned} \tag{4.7}$$

Following the method as discussed in section 3.4 we can derive two equations for the variances,

$$\begin{aligned}\frac{dV_U}{dt} &= \beta_2(p\overline{M}_T + (1-p)\overline{M}_U)) + \mu_M(\overline{M}_U - 2(V_U)) - D_M(2T_U + 4\overline{M}_U V_U - \overline{M}_U^2 - V_U)\\ \frac{dV_T}{dt} &= \beta_2(p\overline{M}_T + (1-p)\overline{M}_U)) + \mu_M(\overline{M}_T - 2(V_T)) - D_M(2T_T + 4\overline{M}_T V_T - \overline{M}_T^2 - V_T).\end{aligned} \tag{4.8}$$

When $p = 1$ and the treatment intensity, u, is variable then assuming that a negative binomial moment closure, which is done to avoid complications of reassigning hosts between treated and untreated classes, we obtain the model system:

$$\begin{aligned}\frac{d\overline{M}}{dt} &= \beta_2\overline{M} - (\mu_M + u)\overline{M} - D_M(V + (\overline{M})^2)\\ \frac{dV}{dt} &= \beta_2\overline{M} + (\mu_M + u)(\overline{M} - 2V) - D_M(\frac{4V^2}{\overline{M}} + 4\overline{M}V - \overline{M}^2 - 3V), \quad \text{for } 0 \le u \le u_{max},\end{aligned} \tag{4.9}$$

it is noted that the control variable u appears in both equations. We make the modelling assumption that fixing the moment closure over the period of treatment is a reasonable simplification.

The cost function (4.5) is extended to include an option on the variance,

$$J(u, \overline{M}, V) = \int_0^{T_f} u^2 + \alpha_1(\overline{M} - M_T)^2 + \alpha_2(V - V_T)^2 dt, \tag{4.10}$$

where V_T is the target value and α_2 is a positive constant. For the variance, V_T is set to be the equilibrium value of V under maximal treatment. We consider control of the variance as a possible scenario of interest by considering the need to maintain a uniform host population in relation to parasite burden for successful agricultural outcomes. The values of α_1 and α_2 are set using the formulae

$$\alpha_1 = \frac{q}{(M_x - M_T)^2}, \qquad \alpha_2 = \frac{1-q}{(V_x - V_T)^2},$$

with V_x being the equivalent to M_x for the variance. The factors q and $(1-q)$ are used as weightings to focus control on either the mean, $q = 1$, or variance, $q = 0$. If it is a combination then the value of q may be set to weight these according to importance.

The Hamiltonian for this problem is given by

$$\begin{aligned}
\mathcal{H}(u, \overline{M}, V, \lambda) &= u^2 + \alpha_1(\overline{M} - M_T)^2 + \alpha_2(V - V_T)^2 \\
&+ \lambda_1 \left(\beta_2 \overline{M} - (\mu_M + u)\overline{M} - D_M(V + (\overline{M})^2)\right) \\
&+ \lambda_2 \left(\beta_2 \overline{M} + (\mu_M + u)(\overline{M} - 2V) - D_M(\frac{4V^2}{\overline{M}} + 4\overline{M}V - \overline{M}^2 - 3V)\right).
\end{aligned} \tag{4.11}$$

The dynamics of the auxiliary variables λ_1 and λ_2 are given by

$$\begin{aligned}
\frac{d\lambda_1}{dt} &= -2\alpha_1(\overline{M} - M_T) - \lambda_1 \left(\beta_2 - \mu_M - u - 2D_M\overline{M}\right) \\
&\quad - \lambda_2 \left(\beta + \mu_M + u - D_M(4V - 2\overline{M} - \frac{4V^2}{\overline{M}^2})\right), \qquad \lambda_1(T_f) = 0 \\
\frac{d\lambda_2}{dt} &= -2\alpha_2(V - V_T) + \lambda_1 D_M - \lambda_2 \left(-2(\mu_M + u) - D_M(\frac{8V}{\overline{M}} + 4\overline{M} - 3)\right), \qquad \lambda_2(T_f) = 0.
\end{aligned} \tag{4.12}$$

The optimality condition for this system is given by

$$\frac{\partial \mathcal{H}(u^*, \overline{M}^*, V^*)}{\partial u} = 2u^* - \lambda_1\overline{M} + \lambda_2(\overline{M} - 2V) = 0. \tag{4.13}$$

which results in the optimal control:

$$u^* = \begin{cases} \frac{\lambda_1\overline{M} + \lambda_2(\overline{M} - 2V)}{2} & \text{if } \frac{\partial \mathcal{H}}{\partial u} = 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0 \\ 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0. \end{cases}$$

Parameter set	Control variable	Parameter	Value
Set 1	Mean	α_1	$\frac{1}{10.25^2}$
		α_2	0
		M_T	9
		V_T	0
Set 2	Variance	α_1	0
		α_2	$\frac{1}{4^2}$
		M_T	0
		V_T	7
Set 3	Mean & Variance	α_1	$\frac{1}{2\times 10.25^2}$
		α_2	$\frac{1}{2\times 4^2}$
		M_T	9
		V_T	7

Table 4.5: Parameter sets for the controls based on the mean, variance or both for the variance inclusive model with cost function defined by (4.10).

As in this model there are three ways that we may measure a treatment's cost:

1) As a function of the mean and treatment cost, $\alpha_2 = 0, \alpha_1 > 0$ or $q = 1$.

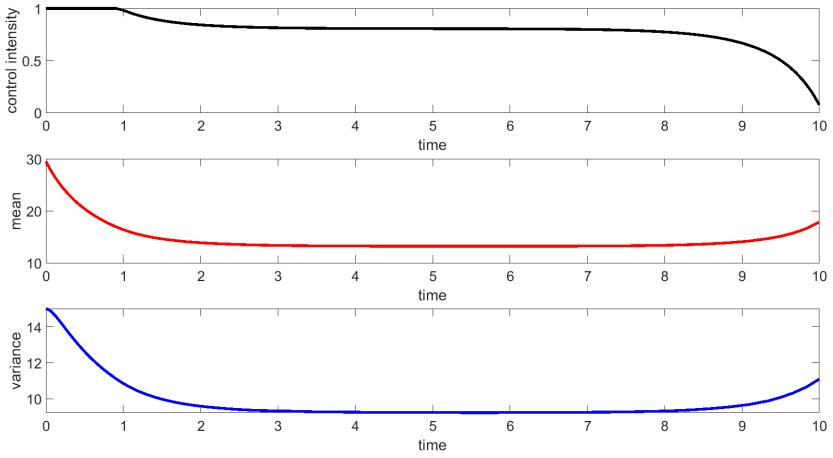
2) As a function of the variance and treatment cost, $\alpha_1 = 0, \alpha_2 > 0$ or $q = 0$.

3) As a function of the the mean, variance and treatment cost, $\alpha_1, \alpha_2 > 0$ or $0 < q < 1$.

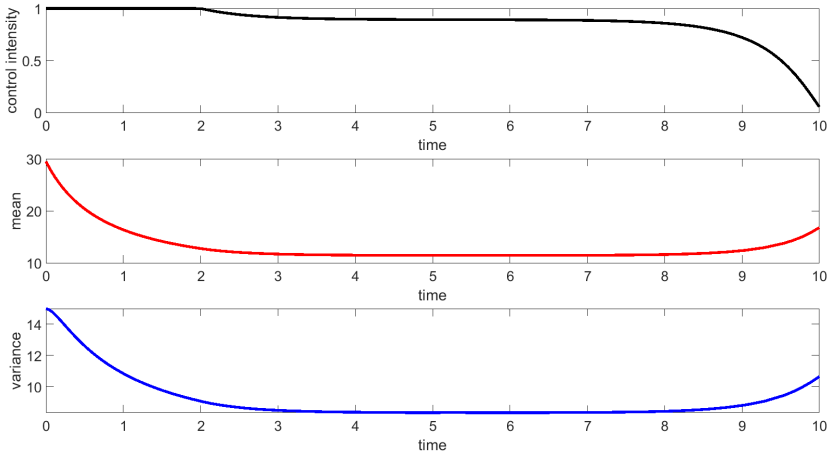
For the examples of control in this thesis on the variance inclusive model, we use the three corresponding parameter sets given in table 4.5. The parameter sets chosen are such that M_T and V_T are approximately equal to the equilibrium values of the dynamic system with the maximal treatment applied. Similar to the mean-field model the values of α_1 and α_2 are selected such that the sum of $\alpha_1(\overline{M} - M_T)^2$ and $\alpha_2(V - V_T)^2$ will be equal to the maximal cost of treatment, i.e 1, when $\overline{M}$ and V are approximately halfway between their treated and untreated equilibrium values.

Figure 4-3 shows three examples of the optimal control and the state dynamics under these controls as determined using the BOCOP software for this model [16].

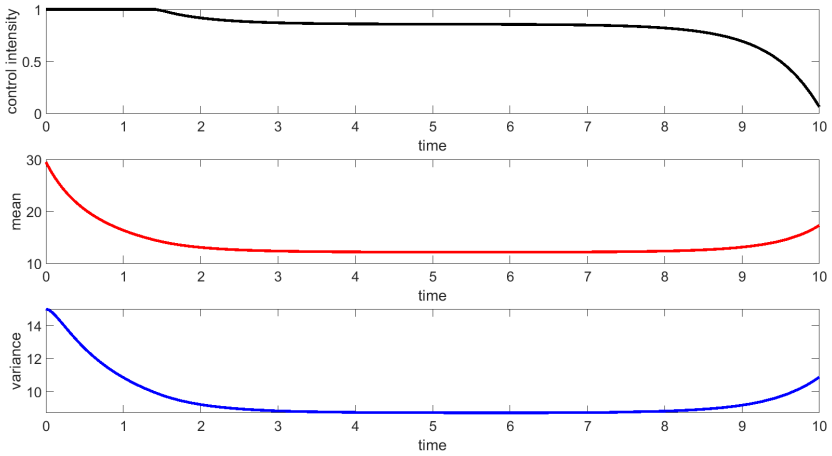
Comparing these controls to the control on the mean-field model shows that the control here behaves similarly, with maximal intensity being applied to begin with before settling at a slightly lower intensity for the majority of the interval and dropping off at the end. The reason for this similarity is that both the mean and the variance respond to the control in very similar ways with each one being reduced as a result. As the aim of the control in each case is to minimise these variables then the control can be applied in very similar ways to achieve the desired results. This is not to say that the inclusion of variance may not be useful in models like this. Comparing the control on the mean against the control on the variance, figure 4-3 sub-figures (a) and (b), we see that the control on the mean has a shorter initial period of maximal intensity. This suggests that the mean is faster to control than the variance in this model.



(a) (a)



(b)



(c)

Figure 4-3: Continuous optimal control on the variance inclusive model for model using cost function (4.10) with parameters (a) Set 1 (b) Set 2 (c) Set 3 from table 4.5.

4.1.5 Application to Isham Model

Taking the negative binomial moment closed Isham model and applying the same type of control as in our models to it would result in the dynamic system:

$$\begin{aligned}\frac{dM}{dt} &= \phi h'(1) - \alpha V(t) - (\mu_M + u(t))M \\ \frac{dV}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M + u)M - 2(\mu_M + u(t))V,\end{aligned} \tag{4.14}$$

with $h(z)$ being the probability generating function of the clump sizes, in this case also of a negative binomial distribution. Using the same cost function as (4.10) leads to the Hamiltonian:

$$\begin{aligned}&\mathcal{H}(u, M, V, \lambda) = u^2 + \alpha_1(M - M_T)^2 + \alpha_2(V - V_T)^2 \\ &+ \lambda_1\left(\phi h'(1) - \alpha V - (\mu_M + u(t))M)\right) \\ &+ \lambda_2\left(\phi(h''(1) + h'(1)) + (\mu_M + u(t))M) - 2(\mu_M + u(t))V\right).\end{aligned} \tag{4.15}$$

The auxiliary variables must satisfy the following system of equation

$$\begin{aligned}\frac{d\lambda_1}{dt} &= -2\alpha_1(M - M_T) + \lambda_1(\mu_M + u(t)) \\ &\quad - \lambda_2(\mu_M + u) \\ \frac{d\lambda_2}{dt} &= -2\alpha_2(V - V_T) + \lambda_1(\alpha) + 2\lambda_2(\mu_M + u(t)).\end{aligned} \tag{4.16}$$

Differentiating the Hamiltonian with respect to u gives the optimality condition

$$\frac{\partial \mathcal{H}(u^*, M^*, V^*)}{\partial u} = 2u^* - \lambda_1 M + \lambda_2(M - 2V) = 0,$$

and hence the optimal control:

$$u^* = \begin{cases} \frac{\lambda_1 M - \lambda_2(M-2V)}{2} & \text{if } \frac{\partial \mathcal{H}}{\partial u} = 0 \\ 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0. \end{cases}$$

The three parameter sets used in examples shown here are given in table 4.6. Figure 4-4 shows the output of the controlled Isham model using the parameters from set 1. It demonstrates that the control acts in a similar way to that employed with models earlier in this chapter (and so the variance only and mean and variance control scenarios are not presented). What the variance and mean only controls did show was that contrary to our variance model is that control will be applied for a longer period when the control is chosen to minimise the mean. One possible explanation is that the Isham model predicts an overdispersed population, with variance greater than mean. Consequently, a control appears more effective when measured by the reduction in the variance than by the reduction in the mean. This theory is supported by figure 4-5, which compares the cost of

Parameter set	Control variable	Parameter	Value
Set 1	Mean	α_1	$\frac{1}{11.5^2}$
		α_2	0
		M_T	1.98
		V_T	0
Set 2	Variance	α_1	0
		α_2	$\frac{1}{20.75^2}$
		M_T	0
		V_T	3.49
Set 3	Mean & Variance	α_1	$\frac{1}{2\times 11.5^2}$
		α_2	$\frac{1}{2\times 20.75^2}$
		M_T	1.98
		V_T	3.49

Table 4.6: Parameter sets for the controls based on the mean, variance or both for the Isham model with cost function defined by (4.10).

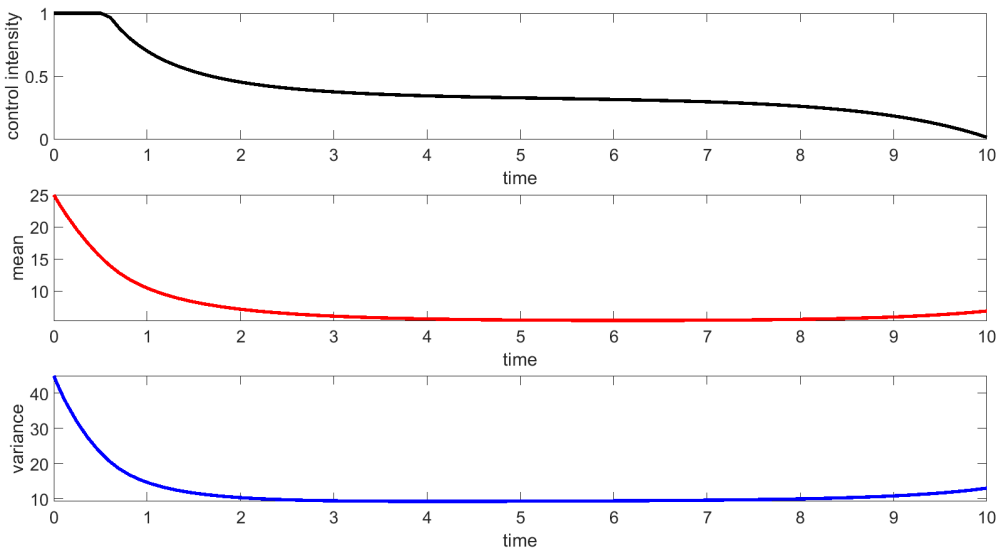
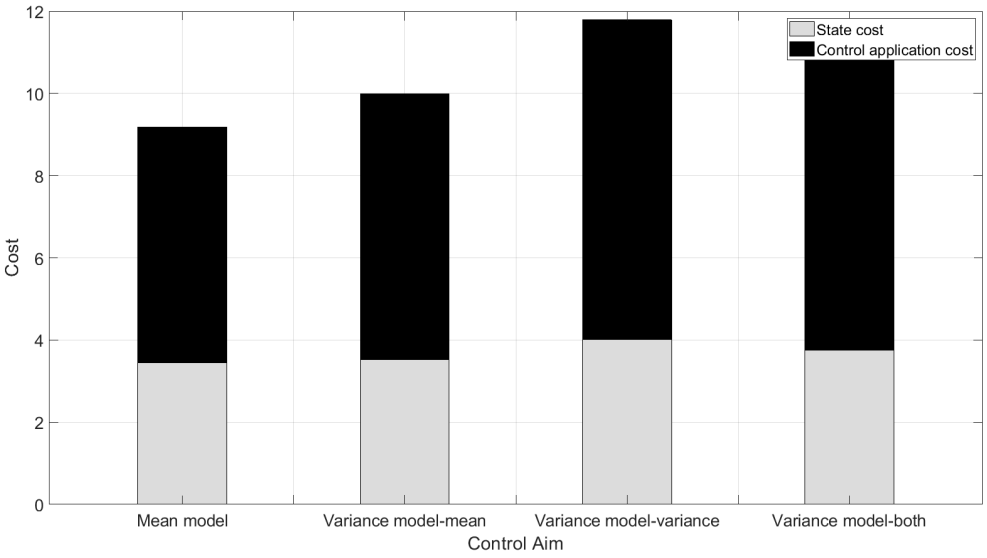


Figure 4-4: Continuous optimal control on the Isham model for model parameters in table 4.3. Optimisation parameters given by parameter set 1, table 4.6, Isham model optimising the mean parasite burden.

the controls implemented for the mean-field, variance, and the Isham models. Sub-figure (b), which shows the cost breakdown of the controls on the Isham models, shows that the cost of control in each case is very similar but that the cost associated with the state variables is much lower when controlling for the variance.

Optimal control is just one approach to infection control. With this in mind, the following section will examine the construction and optimisation of controls which may be more representative of the actual application of control to a host population and the effect on the population dynamics of the parasites within the hosts. Some of these methods are then used in conjunction with more complex cost functions that aim to better represent the aims of control.



(a)

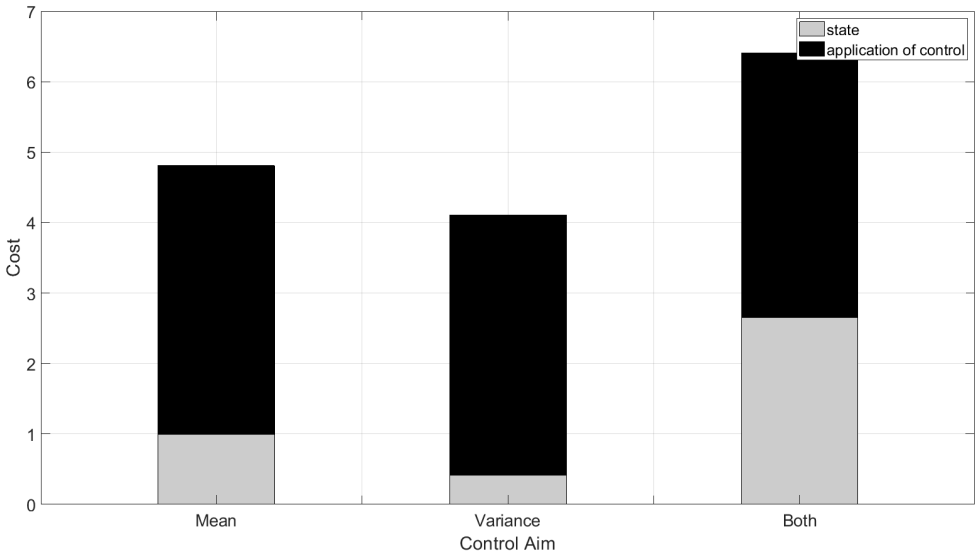


Figure 4-5: (a) Comparison of the cost breakdown of the continuous optimal controls for the basic mean and variance models (b) Comparison of the cost breakdown of the continuous optimal controls on the Isham model.

4.2 Integer Controls

This section begins to look at controls on parasite populations which may either be switched on at a fixed intensity or be switched off. These controls are intended to act in a way that more closely mimics how the treatment would actually be applied, that is as a set dosage that is either given to a host or not. We devised two methods by which control over an interval might be chosen to be representative of practical applications. These are: We devised two methods by which control over an interval might be chosen to be representative of practical applications. These are:

Method 1) To use a set of discrete sub-intervals and choose whether to apply control to each sequentially.

Method 2) To have a control that may change at discrete time points optimised across the whole interval.

We applied the control to the different model structures. These methods are strongly related to one another in that they are both based on discrete sub-intervals of the overall time interval for control. The methods by which these give controls are given below.

Method 1:

1) Initialisation: Set up the dynamics of the model and the cost function, e.g (4.9) and (4.10). Set initial conditions for the mean and variance of the parasite population, (M_0, V_0) equal to the control-free equilibrium values. $(M_0, V_0) = (M^*, V^*)$.

2) Set up a discrete set of time intervals within the overall interval control is being applied over. That is
$$0 = t_1 < t_2 < \cdots < t_N = T_F, \qquad \text{where } t_j - t_{j-1} = \frac{T_F}{N-1}.$$

3) Run the dynamic problem on the j-th interval, $[t_j, t_{j+1}]$, with no control applied and with full control applied from the initial conditions of the interval. If $j = 1$ initial conditions given by $(\overline{M}_0, V_0)$.

4) Calculate the value of the cost function on the interval in both cases. Choose the control option which minimises this.

5) State dynamics in that interval given by trajectory corresponding to chosen treatment.

6) Set initial conditions for $j + 1$-th interval as end conditions of chosen trajectory.

7) If $j \leq N - 1$ set $j = j + 1$ and return to 3), otherwise end.

Method 2:

1) Initialise the system: set the initial cost sufficiently large such that the cost of control will be lower than the cost of current parasite burden and a control strategy will be employed.

2) Set up discrete time steps, $t_0 = \tau_1 < \tau_2 < \cdots < \tau_N = t_F$, with $\tau_{j+1} - \tau_j = \frac{t_F - t_0}{N-1}$, for $j \in \{1, 2, \ldots, N-1\}$.

3) Set up treatment matrix with all admissible controls stored as rows. Matrix will be size $2^{N-1} \times N - 1$.

4) Calculate the population trajectory when the control from row I is applied. Here i denotes the repetition number of the simulation.

5) Calculate the cost of control i.

6) Compare to current cost, if less than update current cost and store i as current treatment.

7) If $i < 2^{N-1}$ set $i = i + 1$ and return to 4), else end.

The first method considers each sub-interval sequentially and chooses whether it is optimal to apply control to that interval purely based on the value of a cost function on that sub-interval. The end state of each sub-interval then becomes the initial condition for the next one. The method results in a single admissible control which is not guaranteed to be the optimal control for the whole interval. By contrast method two checks all admissible controls to see which gives the lowest cost across the whole interval. Before advancing to consider alternate cost functions we first use these methods to establish integer controls using the same cost functions as above, (4.10) and (4.10).

4.2.1 Method 1

Using codes written in Matlab (see appendix C) we applied method one to the mean-field model, the variance inclusive model, and the Isham model using different sub-interval lengths but an overall time interval of length 10. Figure 4-6 shows some examples of this being done, on the mean-field, variance, and Isham models. These controls use the same control functions and parameter sets defined previously for each model which correspond to whichever variable is the focus of the control.

In these figures, the interval length used is equal to 1. Further tests were done using intervals of lengths 0.5 and 0.25, and in the case of the variance and Isham models with cost functions for the other parameter sets given in tables 4.5 and 4.6 that are not shown in these figures. From these figures, and those which have not been included it was noted that we see frequent switching of the control. Due to the way that this method works this tells us that control is not always advantageous when looking at singular intervals. Although not shown here, in performing the same simulation with shorter intervals we were able to see a pattern as a result of the length of the sub-intervals; shorter intervals lead to less control being applied. To see how the length of the intervals affects the cost of control we plotted the cost of control for each interval length for each model and cost function and compared them along with the cost of the continuous control. These are shown for the mean-field model, the variance model with control based on variance, and the Isham model with control based on the mean in figure 4-7.

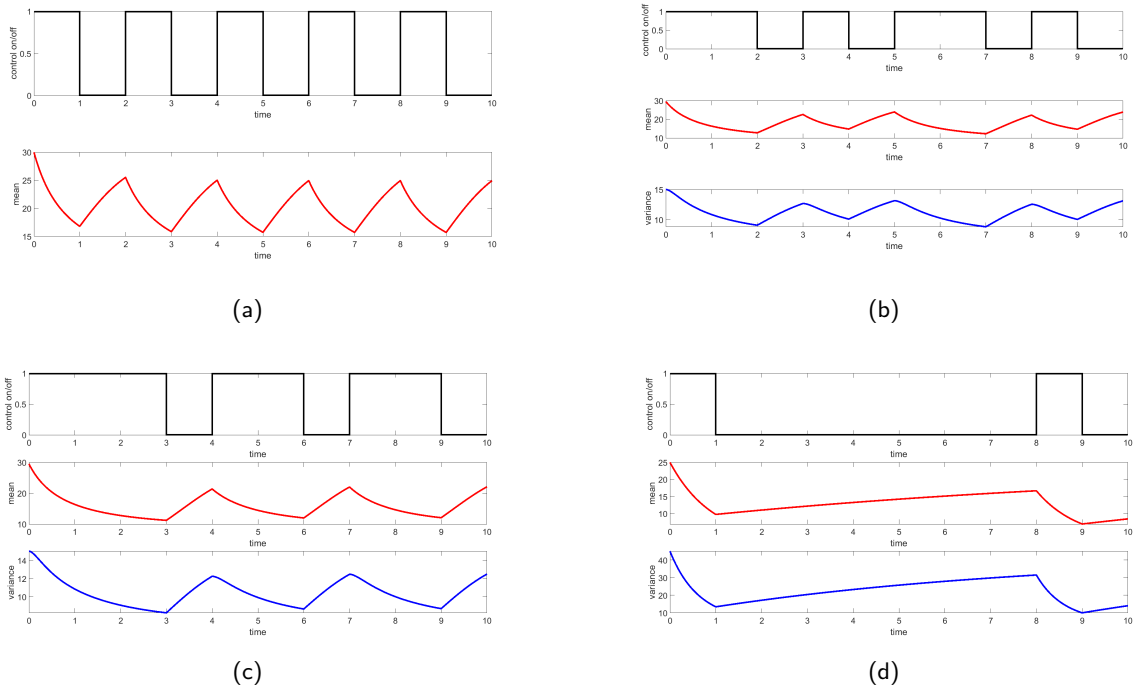
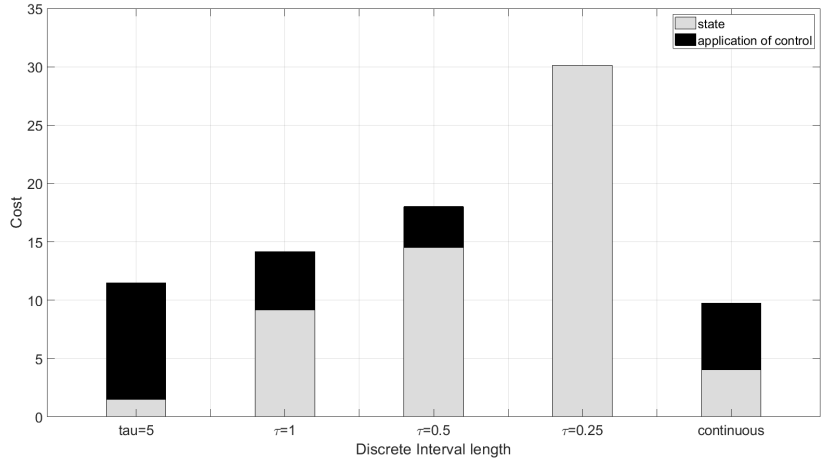


Figure 4-6: Sequential determination of an integer control on (a) the mean-field model (b) Variance inclusive model based controlling the mean (c) Variance inclusive model based controlling the variance (d) Isham model controlling the mean all with sub-interval length 1.

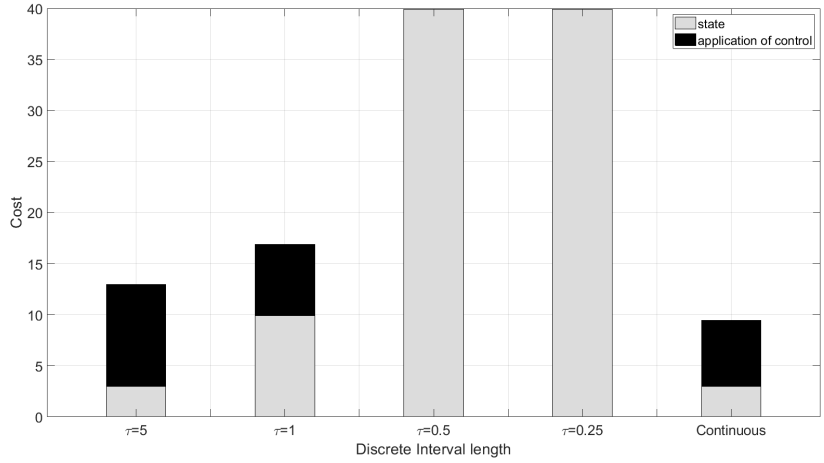
Based solely on the cost of the treatments on the mean-field and variance inclusive models it appears that increasing the sub-interval length is beneficial but the Isham model shows that this is not the case. On the Isham model, increasing the length of the sub-intervals is beneficial to a point but continuing to do so is detrimental to the cost of control. This demonstrates a key feature of optimisation methods; the ability to look ahead.

4.2.2 Method 2

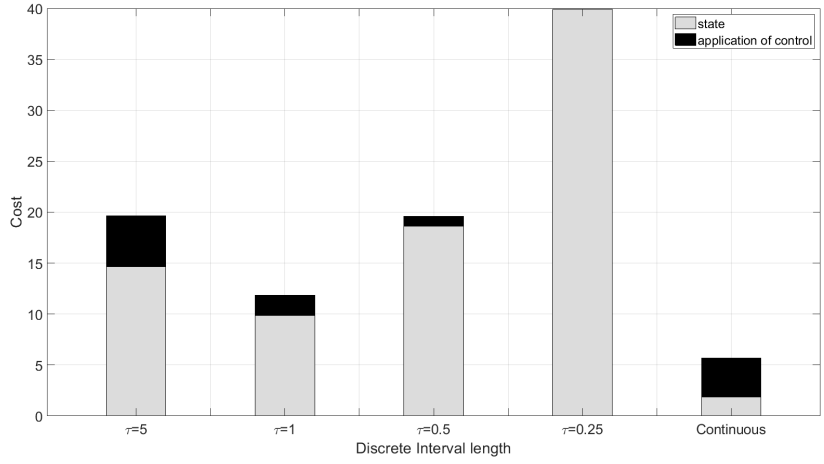
Having used method one to give examples of admissible controls and seen the effect of interval length on the controls and the associated cost we move now onto method two which incorporates the element of optimisation that the results suggested could lead to better integer controls. Owing to the computational cost of the "brute force "method that we are using there are limitations as to how many sub-intervals we can use. Here we limit ourselves to having fewer than fifteen. Given the smallest sub-interval we have used is 0.25 this limits the overall interval length to 3.5. Figure 4-8 shows examples of this for the mean-field model, variance inclusive model, two of the Isham model for some of the cost function parameter sets discussed previously. Further simulations were performed to cover the other cost function parameter sets and interval lengths although they are not shown here.



(a)



(b)



(c)

Figure 4-7: Comparison of the cost of the integer treatments constructed sequentially using different sub-interval lengths, τ, for the (a) Mean-field model (b) Variance model with control based on variance (c) Isham model with control based on mean.

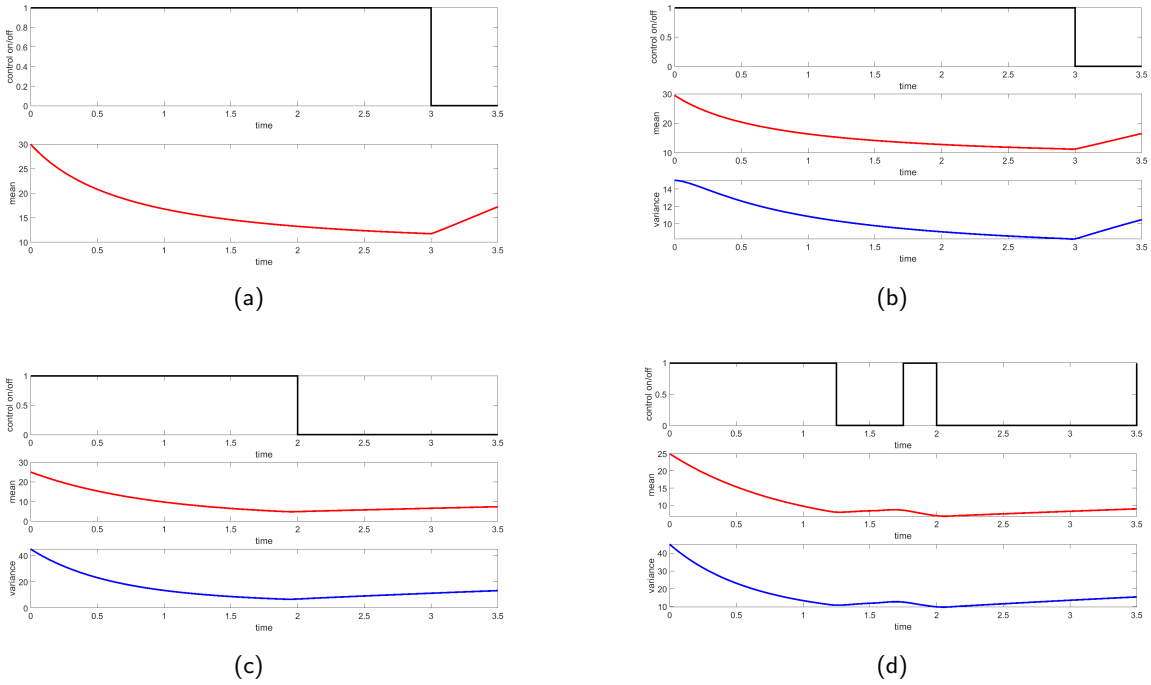


Figure 4-8: Discrete integer optimal control on (a) the mean-field model (sub-interval length 1) (b) Variance model (control on mean, sub-interval length 1) (c) Isham model(control on mean, sub-interval length 1) (d) Isham model (control on mean, sub-interval length 0.25). All model parameters and cost function parameters as previously defined for each model.

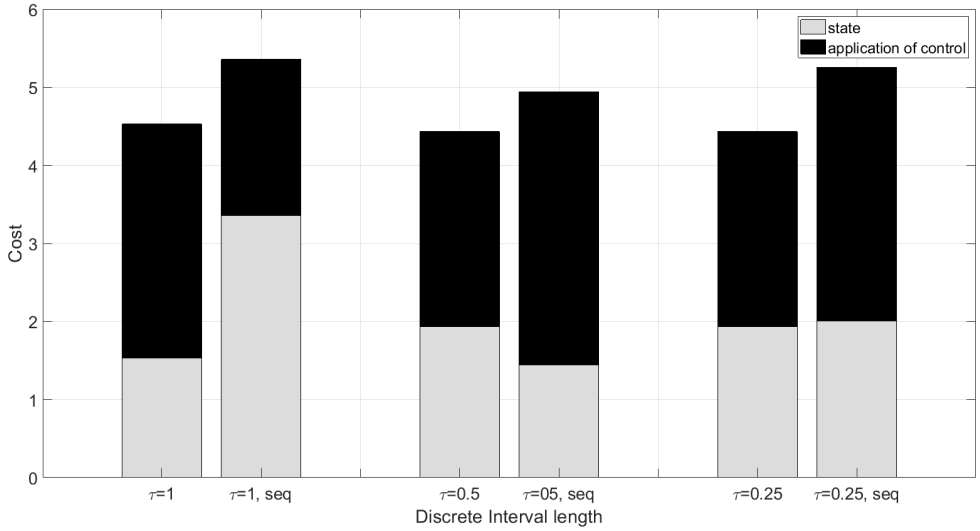


Figure 4-9: Comparison of the cost of the optimal integer controls on the different length discrete sub-intervals and the sequential controls on the same sub-intervals for the mean-field model with model parameters given in table 3.2 and cost function parameters given in table 4.4.

These figures once again share a common property that the optimal control tends to be switched on at the start and maintained until near the end of the overall interval when it is switched off. There were common properties found in many of the models simulated using this method. Both the mean-field model and the variance inclusive model tended to have control applied heavily initially and have the control switched off near the end of the overall interval. For shorter treatment intervals the control was sometimes stopped earlier when compared with a treatment based on the exact same cost function parameters with longer intervals.

The Isham model in some cases varies from this pattern, as shown in figure 4-8 sub-figures (c) and (d). For the shorter interval length, the treatment is switched off then on again before being switched off until the end of the overall interval. This could be due to the Isham model being less resilient, and so slower to return to treated equilibrium, meaning that the control can be switched off for longer periods without seeing as much of an increase in the state variables as the mean-field and variance inclusive models. A lower resilience and either a positive reactivity or simply a smaller magnitude reactivity of the model could demonstrate similar behaviour. The computational cost means that this method is not ideal for a large number of sub-intervals so smaller sub-intervals are more difficult to use. Figure 4-9 shows a cost comparison of the different interval lengths for the control on the mean-field model.

It is interesting that in some cases, such as for the mean model comparison shown, despite the controls being different for the different interval lengths there is not much difference in the overall cost of the controls. What is different is the cost breakdown. This shows that when there are limitations on the control, such as integer values with fixed switching times there could be multiple

controls that give the same lowest value. This raises the question of whether cost functions that take the cost of control and state value and weigh them against one another is the best way to implement control. This is a motivating factor in considering the methodology presented in the next section.

4.3 Control as a Switching System

This method of determining control is another way in which we can determine an integer control. It differs from the two methods presented above in that the benefit of turning on or off control is not calculated by the value of a cost function over any length interval. It is instead a form of switched system. A switched system involves having the dynamics of the model defined when certain conditions are met, once these conditions are no longer met the system switches to an alternate set of dynamics [59]. It may be that the system switches back again once the original conditions are met once more or the system may involve a "buffer zone", such that the dynamics within the buffer zone depend on where the system was previously. This can help prevent situations where the dynamics switch rapidly between the two systems. Figure 4-10 shows the difference between these two systems. If the state enters the buffer zone from the other side it remains under the dynamics of system two until the criteria of switch two are reached. The examples that we present here use a method similar to that with the buffer zone. A set condition switches treatment on and an alternate condition switches it back off. For each of the models we use the same model parameters as previously, shown in table 4.1 for the mean-field model, table 4.2 for the variance inclusive model and table 4.3 for the Isham model.

With the switching systems that we have here the basis of the model dynamics both during and in the absence of treatment are the same. The switching system simply operates on the control parameter, u.

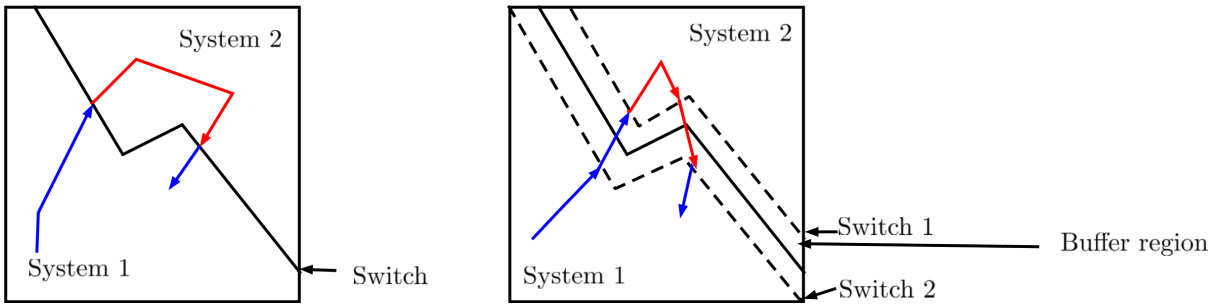


Figure 4-10: Left: Switching system with a single switch. Right: Switching system with two switches and a buffer zone.

4.3.1 Application to the Mean-Field Model

To apply this method to the mean-field model we set up the switched dynamic system:

$$\begin{aligned}\frac{dM}{dt} &= \beta_2 M - (\mu_M + u_{min})M - D_M M^2, \text{ if } f_1(M(t), u(t-\delta t)) = 1 \\ \frac{dM}{dt} &= \beta_2 M - (\mu_M + u_{max})M - D_M M^2, \text{ if } f_1(M(t), u(t-\delta t)) = 0,\end{aligned} \tag{4.17}$$

where $f_1(M(t), u(t-\delta t))$ is a function which maps every point in the domain of M to a value of either zero or one but also takes into account the value of the control immediately preceding the current time. To show how a control derived using this method differs from the controls shown previously we set up a system that would turn on control at a set value of the mean, M_x. To switch control back off we considered two potential options; switch off after a set time to mimic a course of treatment or switch off when the treatment is no longer effective enough, for example, if

$$\frac{dM}{dt} > -c.$$

The simpler scenario may be written as

$$f_1(\overline{M}, u(t-\delta t)) = \begin{cases} 0 & \text{if } u(t-\delta t) = 0 \text{ and } \overline{M} \geq M_x \\ 1 & \text{if } u(t-\delta t) = 0 \text{ and } \overline{M} < M_x \\ 0 & \text{if } u(t-\delta t) = 1 \text{ and } \tau \geq T_x \\ 1 & \text{otherwise,} \end{cases} \tag{4.18}$$

where τ is the time since the current treatment began and T_x is the length of time a single treatment is prescribed to last for. The second scenario has a switch function given by:

$$f_1(\overline{M}, u(t-\delta t)) = \begin{cases} 1 & \text{if } u(t-\delta t) = u_{min} \text{ and } \overline{M}(t) < M_x \\ 0 & \text{if } u(t-\delta t) = u_{min} \text{ and } \overline{M}(t) \geq M_x \\ 0 & \text{if } u(t-\delta t) = u_{max} \text{ and } \frac{d\overline{M}}{dt} \leq -c \\ 1 & \text{if } u(t-\delta t) = u_{max} \text{ and } \frac{d\overline{M}}{dt} > -c. \end{cases} \tag{4.19}$$

It is possible with this switching function that treatment could be switched on due to the mean exceeding M_x and immediately switched off due to $\frac{d\overline{M}}{dt} > c$. In cases where this occurs a choice must be made as to whether the current burden or the ability of the control to reduce the burden is more important.

Figure 4-11 sub-figures (a) and (c), show two examples of the controls given by these methods, one with a treatment switched off by time (switch 1, switched off after $\tau = 0.45$ units) and the other by an assessment of the efficacy of the treatment (switch 2, switched off by $\frac{dM}{dt} > -1$). Sub-figures (b) and (d) show a comparison of the cost of these controls, using the cost function

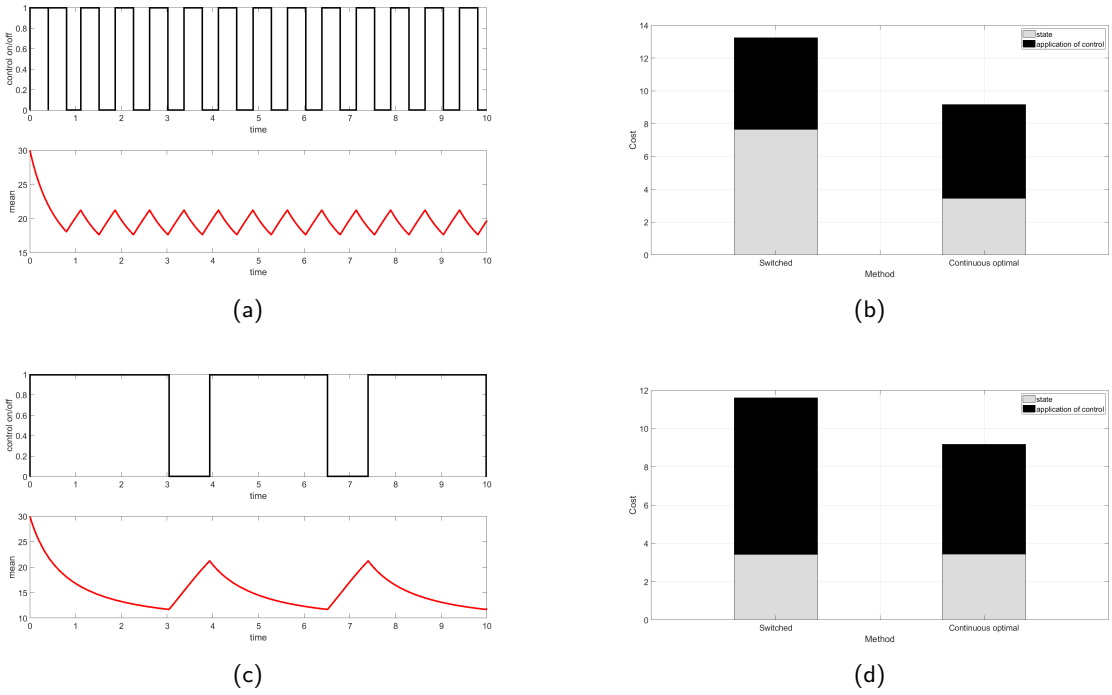


Figure 4-11: (a) Treatment applied to the mean-field model as a result of using switch 1 (b) Comparison of the cost of using switch 1 on the mean-field model against the continuous control (c) Treatment applied to the mean-field model as a result of using switch 2 (d) Comparison of the cost of using switch 2 on the mean-field model against the continuous control.

from (4.5), against the cost of continuous control. This comparison is made due to the use of the switch on being set at $M_x = 21.24$ which is the value of the mean which we used to set our cost function parameter α_1 in that case.

4.3.2 Application to the Variance Inclusive Model

For the variance inclusive model, the switching system is very similar to that of the mean-field model. The main difference is that there are more potential switching functions available. Using the dynamic systems already defined for the variance inclusive model then the switching system may be written as:

$$\begin{aligned} \frac{dM}{dt} &= \beta_2 M - (\mu_M + u_{min})M - D_M(V + M^2) \text{ when } f_1(M, V, u(t - \delta t)) = 1 \\ \frac{dV}{dt} &= \beta_2 M + (\mu_M + u_{min})(M - 2V) - D_M(\frac{4V^2}{M} + 4MV - M^2 - 3V) \end{aligned} \tag{4.20}$$

and

$$\begin{aligned} \frac{dM}{dt} &= \beta_2 M - (\mu_M + u_{max})M - D_M(V + M^2) \text{ when } f_1(M, V, u(t - \delta t)) = 0 \\ \frac{dV}{dt} &= \beta_2 M + (\mu_M + u_{max})(M - 2V) - D_M(\frac{4V^2}{M} + 4MV - M^2 - 3V) \end{aligned} \tag{4.21}$$

The switches based on the mean are the same as equations (4.19) and (4.18) triggers based on the variance, the rate of change of the variance may be written as:

$$f_1(M,V,u(t-\delta t)) = \begin{cases} 0 & \text{if } u(t-\delta t) = u_{min} \text{ and } V \geq V_x \\ 1 & \text{if } u(t-\delta t) = u_{min} \text{ and } V < V_x \\ 10 & \text{if } u(t-\delta t) = u_{max} \text{ and } u(t-\tau) = u_{min}, \text{ for some } 0 < \tau < T_x \\ 1 & \text{otherwise,} \end{cases} \tag{4.22}$$

$$f_1(M,V,u(t-\delta t)) = \begin{cases} 0 & \text{if } u(t-\delta t) = u_{min} \text{ and } V \geq V_x \\ 1 & \text{if } u(t-\delta t) = u_{min} \text{ and } V < V_x \\ 1 & \text{if } u(t-\delta t) = u_{max} \text{ and } \frac{dV}{dt} > -c \\ 0 & \text{otherwise.} \end{cases} \tag{4.23}$$

It is also possible that the switch functions could account for one variable causing the activation of treatment and the other the deactivation or even a combination of both variable causing the switch. Some examples of this are shown in figure 4-12. The switch mechanisms in use are:

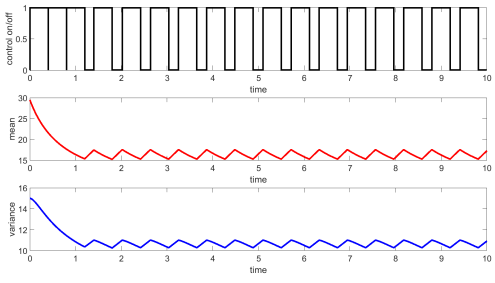
- Switch 3 - Equation (4.22), with $V_x = 11$ and $\tau = 0.4$
- Switch 4 - Equation (4.23), with $V_x = 11$ and $c = -1$.

When the variance model employed a switching system based on the mean and the rate of change of the mean the results were very similar to those seen on the mean-field model.

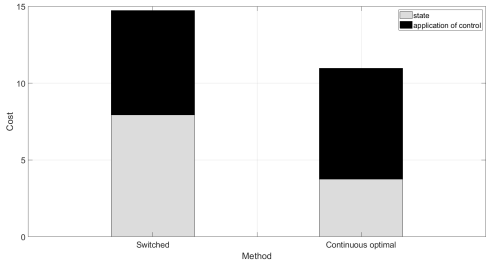
4.3.3 Application to the Isham Model

In this section, we will use the same method applied to the Isham model. However, we will also show how determining control using a method such as this can allow more complex criteria to be used in control decisions. First, we look at using switching functions which are exactly the same as for the variance inclusive model switches shown above but used in conjunction with the following dynamics:

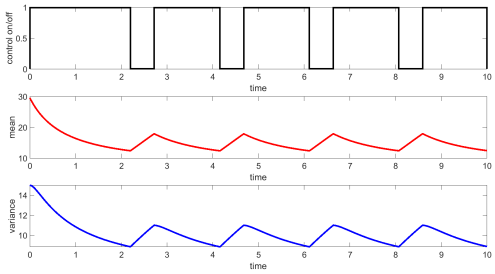
$$\begin{aligned} \frac{dM}{dt} &= \phi h'(1) - \alpha V(t) - (\mu_M + u_{min})M \\ \frac{dV(t)}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M + u_{min})M - 2(\mu_M + u_{min})V(t), \text{ if } f_1(M,V,u(t-\delta t)) = 1, \end{aligned} \tag{4.24}$$



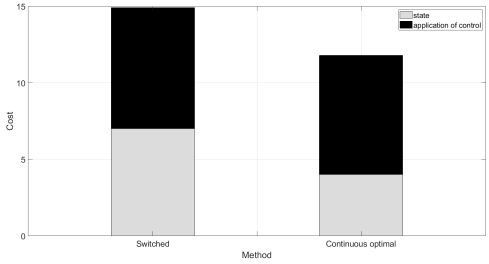
(a)



(b)



(c)



(d)

Figure 4-12: (a) Treatment applied to the variance inclusive model using switch 3 (b) Comparison of the cost of using switch 3 against the continuous control (c)Treatment applied to the variance inclusive model using switch 4 (b) Comparison of the cost of using switch 4 against the continuous control

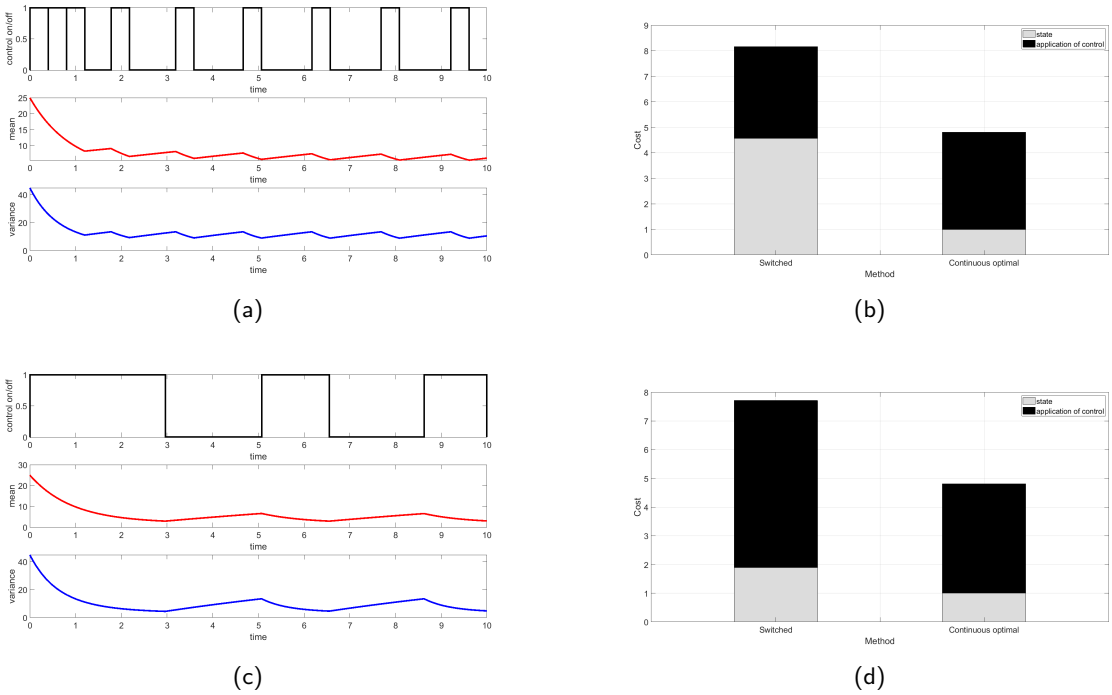


Figure 4-13: (a) Treatment applied to the Isham model using switch 5 (b) Comparison of the cost of using switch 5 against the continuous control (c) Treatment applied to the Isham model using switch 6 (b) Comparison of the cost of using switch 6 against the continuous control

and

$$\begin{aligned}\frac{dM}{dt} &= \phi h'(1) - \alpha V(t) - (\mu_M + u_{max})M \\ \frac{dV(t)}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M + u_{max})M - 2(\mu_M + u_{max})V(t), \text{ if } f_1(M, V, u(t-\delta t)) = 0,\end{aligned} \tag{4.25}$$

Examples of these some of these controls can be seen in the figures 4-13 and 4-14. In this figure the switches used are as follows:

- Switch 5: Equation (4.18) - switched on when M exceeds 13.49 and switched off after a time f $\tau = 0.4$
- Switch 6: Equation (4.19) - switched on when M exceeds 13.49 and switched off when $\frac{dM}{dt} > -1$.
- Switch 7: Equation (4.22) - switched on when V exceeds 24.245 and switched off after a time f $\tau = 0.4$
- Switch 8: Equation (4.23) - switched on when M exceeds 24.245 and switched off when $\frac{dV}{dt} > -1$.

These figures show how the different control strategies can affect the parasite dynamics in a range of ways. Controls that target the mean, in this case, are employed more frequently than those

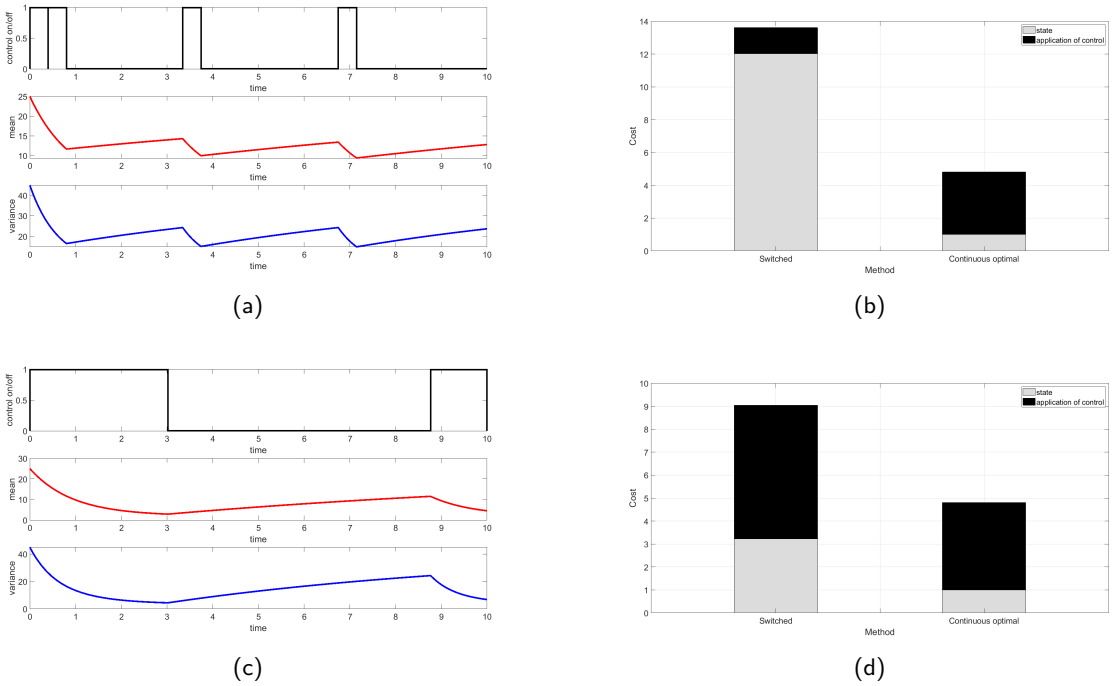


Figure 4-14: (a) Treatment applied to the Isham model using switch 7 (b) Comparison of the cost of using switch 7 against the continuous control (c) Treatment applied to the Isham model using switch 8 (b) Comparison of the cost of using switch 8 against the continuous control

which target the variance. This is as a result of the mean being much smaller than the variance, which means that a much smaller increase is needed before the treatment is triggered once more. In addition, the rate of increase of the mean, while in totality lower than the rate of increase of the variance; is greater as a proportion of the overall size of the mean compared with the rate of increase of the variance as a proportion of the size of the variance.

While the difference that we see in these controls in comparison to the mean-field and variance inclusive models is interesting perhaps the most vital property that we show with the Isham model is in our alternative switching functions. Perhaps the most obvious idea for a switch mechanism is to turn treatment on when it is believed that some hosts are burdened to such an extent that they would begin to suffer consequences. For example, if some percentage of hosts were predicted to be burdened above a set level. Although the ultimate goal may be to treat only those hosts to avoid unnecessary treatment we will first focus on identifying when this occurs.

The Isham model is important here as the mean and the variance fit well with a negative binomial distribution. If we continue to assume that the parasite populations described by the model have a negative binomial distribution we may use the mean and the variance to determine the parameters of this distribution at any time using that for a negative binomially distributed random

variable $M \sim NB(p, r)$ the mean and the variance in the size of this variable will be given by

$$\overline{M} = \frac{pr}{1-p} \text{ and } V = \frac{pr}{(1-p)^2}. \tag{4.26}$$

By rearranging we may use the mean and variance from the model to give the parameters

$$p = \frac{\overline{M}}{V} \text{ and } r = \frac{\overline{M}^2}{V - M}. \tag{4.27}$$

Now for the negative binomial function, we may use that the cumulative distribution function is given by

$$\mathbb{P}(M \leq m) = I_p(m+1, r) = \frac{B(p; m+1, r)}{B(m+1, r)}, \tag{4.28}$$

where $I_p(m+1, r)$ is the regularised incomplete Beta function. While analytically this method is complex, it is quite simple to simulate using Matlab. So we may set up our switch function as

$$f_1(\overline{M}, V, u(t-\delta t)) = \begin{cases} 1 & \text{if } u(t-\delta t) = u_{min} \text{ and } I_p(m, r) < p_c \\ 0 & \text{if } u(t-\delta t) = u_{min} \text{ and } I_p(m, r) \geq p_c \\ 0 & \text{if } u(t-\delta t) = u_{max} \text{ and } u(t-\tau) = u_{min}, \text{ for some } 0 < \tau < T_x \\ 1 & \text{otherwise.} \end{cases} \tag{4.29}$$

Here m denotes the burden at which we aim to keep the majority below and p_c the proportion of hosts that must be infected above this level for us to begin treatment. An example of this is shown in figure 4-15.

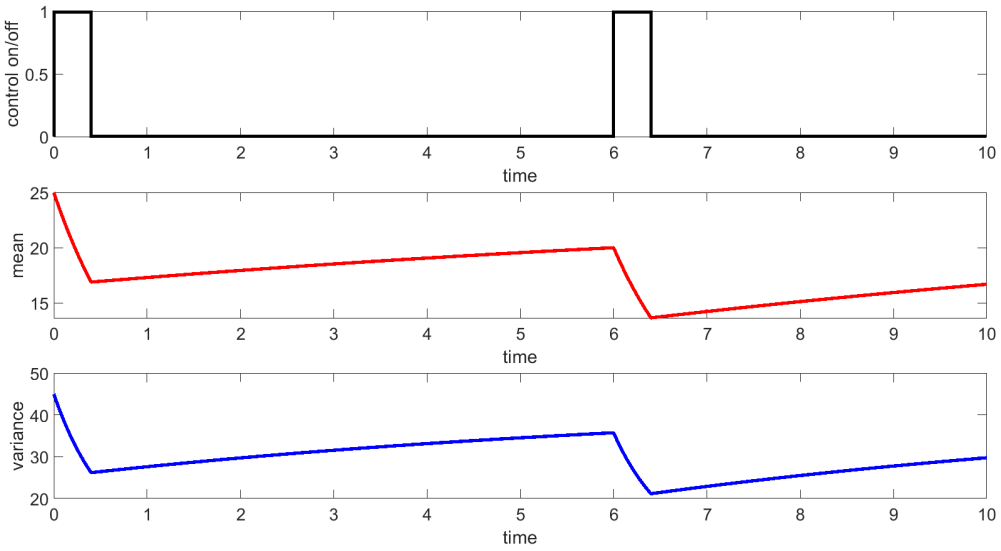


Figure 4-15: Example of a treatment on the Isham model which is determined by a switching system. The control is switched on when the proportion of hosts predicted to be infected with more than 30 parasites exceeds 0.05.

4.4 Discussion

The key findings that we have discussed in this chapter are:

1. The underlying dynamics of treatment require careful consideration, it may be necessary to return to the individual scale models to determine how a treatment affects the underlying dynamics.

2. Where treatment is applied to the whole population and treatments have a similar effect on the mean and variance, for example, both are reduced, and the goal of treatment is to utilise this effect, whether it is on the mean or the variance, then the optimal controls in each case will be markedly similar.

3. With more information on the distribution however the variance may be useful in creating new measures of the effectiveness of a treatment or even creating a switching system that can account for a more specific and potentially complex property of the distribution of parasites.

In the following chapters, we expand our models further such that the models account for potential resistance to treatment.

Chapter 5

Resistant Parasite Models

In the treatment of macroparasitic diseases, there have been serious problems with parasites developing resistance to the drugs that are used to treat them. The problem is so widespread that the American Food and Drug Administration (FDA) has conducted studies and published advisory strategies to try and minimise the problem [2]. As such any control should account for the potential increase in parasite resistance to treatments. This chapter establishes and analyses model structures which are used in the following chapter to explore the impact of control when resistance to control interventions are present. We do this in a very simplistic way by creating a system that couples the dynamics of two parasite populations simply by increasing the density dependent death of each population. We assume that the demographic parameters (intrinsic birth rate, per capita mortality rate, etc.) for both populations are the same; the only heterogeneity occurs in the density dependent mortality rates. For continuity with the following Chapter, we denote the two populations susceptible and resistant but acknowledge that these terms only become relevant in Chapter 6 where the impact of control demonstrates the impact of resistant parasites.

5.1 Haploid Models

We begin first with a model that assumes that each parasite has haploid genetics and its resistance status, which is resistant or susceptible, is based on a single allele. These models are formulated in a very similar way to the model we derived in Chapter 3. As a result, we begin by considering an individual simulation model.

In the individual simulation we have a population of N hosts indexed by the numbers $i = 1, 2, \cdots, N$. Within the i-th host we denote the size of the two different genotypes of parasites by $M_{i,R}$ and $M_{i,S}$ for resistant and susceptible respectively. The parasites within the host produce eggs which then pass out of the host and into a shared larval pasture, which has resistant and susceptible populations of size L_R and L_S. From here the parasites may be picked up by any of the hosts where they mature and continue the reproductive cycle. Throughout this cycle, the processes which dictate the birth, pick up and death of the parasites are assumed to be Poisson process and are given by:

- Birth: This is considered to be an asexual process with all parasites giving birth at rate β. The offspring are born into the corresponding larval population.

- Pick up: Individual larval parasites are picked up by a host at rate ϕ, with each host equally likely to pick up the parasites at any time.

- Maturation: The probability that a larval parasite matures once it is ingested by a host is given by ρ, any that do not mature die.

- Larval death: Parasites in the larval stage may also die of natural causes at a constant rate μ_L.

- Mature parasite death: Parasites within a host may die of natural causes as a constant rate μ_M.

- Density dependent mortality: Acknowledging the simplicity of this assumption, we take the per capita density dependent death rate of susceptible parasites as

$$D_M([M_{i,S}] + (1-\psi)[M_{i,R}]),$$

where ψ denotes the probability that the density dependent mortality arising from the cross-interaction of susceptible and resistant parasite populations results in mortality for a susceptible parasite. The corresponding density dependent mortality for resistant parasites is

$$D_M([M_{i,S}] + \psi[M_{i,R}]).$$

Taking these simple processes and using the model construction process outlined in Chapter 3, gives the following equations which denote the rate of change of the two larval populations, and the size of the susceptible and resistant populations within the i-th host:

$$\begin{aligned}
\frac{dL_S}{dt} &= \beta \sum_{j=1}^{N} [M_{j,S}] - \mu_L L_S - N\phi L_S \\
\frac{dL_R}{dt} &= \beta \sum_{j=1}^{N} [M_{j,R}] - \mu_L L_R - N\phi L_R \\
\frac{d[M_{i,S}]}{dt} &= \rho\phi L_S - \mu_M [M_{i,S}] - D_M([M_{i,S}]^2 + (1-\psi)[M_{i,S}][M_{i,R}]) \\
\frac{d[M_{i,R}]}{dt} &= \rho\phi L_R - \mu_M [M_{i,R}] - D_M([M_{i,R}]^2 + \psi[M_{i,S}][M_{i,R}]).
\end{aligned} \tag{5.1}$$

Taking the mean across all hosts leads to the system

$$\begin{aligned}
\frac{dL_S}{dt} &= \beta N \overline{M}_S - \mu_L L_S - N\phi L_S \\
\frac{dL_R}{dt} &= \beta N \overline{M}_R - \mu_L L_R - N\phi L_R \\
\frac{dM_S}{dt} &= \rho\phi L_S - \mu_M M_S - D_M(M_S^2 + V_S + (1-\psi)(M_S M_R + C_{S,R})) \\
\frac{dM_R}{dt} &= \rho\phi L_R - \mu_M M_R - D_M(M_R^2 + V_R + \psi(M_S M_R + C_{S,R})),
\end{aligned} \tag{5.2}$$

where $C_{S,R}$ denotes the covariance between the susceptible and resistant parasite populations. We consider two separate cases following the previous chapter.

In this model we present a similar case but must also account for the covariance between the two genotypes in the sub-populations, $C_{S,R}$. The two models presented in this section are such that:

1. The variances of the genotype sub-populations are equal to zero, $V_S = V_R = 0$. By extension of this $[M_{i,S}] = \overline{M}_S$ and $[M_{i,R}] = \overline{M}_R$, which means that the covariance $C_{S,R} = \frac{1}{N}\sum_{i=1}^{N}([M_{i,S}] - \overline{M}_S)([M_{i,R}] - \overline{M}_R)$ must also be equal to zero.
2. For simplicity it is assumed that $C_{S,R} = 0$. While realistically this will not be the case a lack of data surrounding the actual covariance prevents a more accurate assumption being made. The variances of the genotype sub-populations of the host are modelled directly.

5.1.1 Two Genotype Mean-Field Model

If we first formulate the simpler model then by setting all variances and the covariance equal to zero the model becomes:

$$\begin{aligned}
\frac{dL_S}{dt} &= \beta N \overline{M}_S - \mu_L L_S - N\phi L_S \\
\frac{dL_R}{dt} &= \beta N \overline{M}_R - \mu_L L_R - N\phi L_R \\
\frac{d\overline{M}_S}{dt} &= \rho\phi L_S - \mu_M \overline{M}_S - D_M(\overline{M}_S^2 + (1-\psi)\overline{M}_S\overline{M}_R) \\
\frac{d\overline{M}_R}{dt} &= \rho\phi L_R - \mu_M \overline{M}_R - D_M(\overline{M}_R^2 + \psi\overline{M}_S\overline{M}_R).
\end{aligned} \tag{5.3}$$

Although the simulation index has been dropped, the mean is then taken across an infinite number of populations with the variance between simulations assumed to be zero. This assumption, combined with the simplified notation used here means that the model is still written as in equation 5.3.

Assuming that both the larval populations are at equilibrium at all times the model may be simplified by using that

$$L_S = \frac{\beta N \overline{M}_S}{\mu_L + N\phi} \text{ and } L_R = \frac{\beta N \overline{M}_R}{\mu_L + N\phi}.$$

Substituting this into the equations for the means gives

$$\begin{aligned}\frac{dM_S}{dt} &= \beta_2 M_S - \mu_M M_S - D_M(M_S^2 + (1-\psi)M_S M_R)\\ \frac{dM_R}{dt} &= \beta_2 M_R - \mu_M M_R - D_M(M_R^2 + \psi M_S M_R),\end{aligned} \tag{5.4}$$

where $\beta_2 = \frac{\rho\phi\beta N}{\mu_L + N\phi}$.

5.1.2 Two Genotype Variance Inclusive Model

For the second option we assume the variances , V_S and V_R, are non-zero. Following the methodology described in Chapter 3, we are able to derive the following pair of coupled differential equations to describe the evolution of variances, by assuming a negative binomial distribution for the moment closure:

$$\begin{aligned}\frac{dM_S}{dt} &= \beta_2 M_S - \mu_M M_S - D_M(M_S^2 + (1-\psi)M_S M_R + V_S)\\ \frac{dM_R}{dt} &= \beta_2 M_R - \mu_M M_R\\ &\quad - D_M(M_R^2 + \psi M_R M_S + V_r)\\ \frac{dV_S}{dt} &= \beta_2 M_S - \mu_M(2V_S - M_S)\\ &\quad - D_M(\frac{4V_S^2}{M_S} + 4M_S V_S - M_S^2 - 3V_S + (1-\psi)(2M_R V_S - M_R M_S))\\ \frac{dV_R}{dt} &= \beta_2 M_R - \mu_M(2V_R - M_R)\\ &\quad - D_M(\frac{4V_R^2}{M_R} + 4M_R V_R - M_R^2 - 3V_R + \psi(2M_S V_R - M_S M_R)),\end{aligned} \tag{5.5}$$

where M_S, M_R, V_S and V_R denote the mean and variance that has been averaged across all simulations and $\beta_2 = \frac{\rho\phi\beta N}{\mu_L + N\phi}$.

5.2 Haploid Model Analysis

To reiterate, this model form is simply being introduced so that the impact of control on a parasite population that divides into two control responses can be explored in the following chapter. Not withstanding that, we now undertake some standard equilibrium analysis to highlight key model behaviours.

5.2.1 Two Genotype Mean-Field Model Analysis

If the model shown in (5.4) is considered, it can be seen that if either M_S, or M_R is set to zero the model is essentially the basic mean-field model from (3.8). As a result there are two potential equilibria where only a single genotype survives. These equilibria will be given by

$$(M_S^*, M_R^*) = (\frac{\beta_2 - \mu_M}{D_M}, 0) \text{ or } (M_S^*, M_R^*) = (0, \frac{\beta_2 - \mu_M}{D_M}), \tag{5.6}$$

both of which are only relevant if $\beta_2 > \mu_M$. From the analysis of (3.8) it is known that the model will have an equilibrium at

$$(M_S^*, M_R^*) = (0, 0),$$

which is easily confirmed by substituting these values into (5.4). Of greater interest is the potential equilibria in which both genotypes survive. If both ODEs from (5.4) are set equal to zero it can be shown that

$$\begin{aligned}
M_S(\beta_2 - \mu_M - D_M(M_S + (1-\psi_1)M_R)) &= 0 \\
M_R(\beta_2 - \mu_M - D_M(M_R + \psi_1 M_S)) &= 0 \\
&\Longrightarrow \\
M_S + (1-\psi)M_R &= \frac{\beta_2 - \mu_M}{D_M} \\
M_R + \psi M_S &= \frac{\beta_2 - \mu_M}{D_M} \\
&\Longrightarrow \\
M_S(1 - (1-\psi)\psi) &= (1 - (1-\psi))\frac{\beta_2 - \mu_M}{D_M} \\
M_R(1 - \psi(1-\psi)) &= (1-\psi)\frac{\beta_2 - \mu_M}{D_M} \\
\Longrightarrow\ M_S &= \frac{\psi(\beta_2 - \mu_M)}{D_M(1 - \psi + \psi^2)} \\
M_R &= \frac{(1-\psi)(\beta_2 - \mu_M)}{D_M(1 - \psi + \psi^2)},
\end{aligned}$$

which gives that the equilibria where both genotypes survive is given by:

$$(M_S^*, M_R^*) = \left(\frac{\psi(\beta_2 - \mu_M)}{D_M(1 - \psi + \psi^2)}, \frac{(1-\psi)(\beta_2 - \mu_M)}{D_M(1 - \psi + \psi^2)}\right).$$

The condition for this equilibrium to be such that the values of M_S^* and M_R^* are positive is that $\beta_2 > \mu_M$. This is due to the original condition that $0 \leq \psi, (1-\psi) \leq 1$ which must be imposed as they denote a probability.

As discussed in the analysis of (3.8) the trivial equilibrium is stable only if $\beta_2 < \mu_M$ and all other equilibria become negative. However before determining the stability of the other equilibria further

investigation is required. For the model the Jacobian may be derived as

$$J = \begin{bmatrix} \beta_2 - \mu_M - D_M(2M_S + (1-\psi)M_R) & -D_M(1-\psi)M_S \\ -D_M\psi M_R & \beta_2 - \mu_M - D_M(2M_R + \psi M_S) \end{bmatrix}. \tag{5.7}$$

For the equilibria where only the susceptible genotype survives this will be given by

$$J\left(\frac{\beta_2 - \mu_M}{D_M}, 0\right) = \begin{bmatrix} \mu_M - \beta_2 & (1-\psi)(\mu_M - \beta_2) \\ 0 & (\beta_2 - \mu_M)(1-\psi), \end{bmatrix}$$

which has eigenvalues

$$\lambda_1 = (\beta_2 - \mu_M)(1-\psi) \text{ and } \lambda_2 = \mu_M - \beta_2.$$

This shows that this will only be stable if $\beta_2 > \mu_M$ and $1 - \psi < 0$, which violates the original condition on ψ. In a similar manner the Jacobian where only the resistant parasites survive has Jacobian given by

$$J\left(0, \frac{\beta_2 - \mu_M}{D_M}\right) = \begin{bmatrix} (\beta_2 - \mu_M)\psi & 0 \\ \psi(\mu_M - \beta_2) & \mu_M - \beta_2, \end{bmatrix}$$

and eigenvalues

$$\lambda_1 = (\beta_2 - \mu_M)\psi \text{ and } \lambda_2 = \mu_M - \beta_2.$$

For stability this would require that $\psi < 0$ which would also violate the original conditions of the model. Moving now onto the equilibria wherein both genotypes coexist the Jacobian will be given by

$$J(M_S^*, M_R^*) = \begin{bmatrix} (\beta_2 - \mu_M)(1 - \frac{1+\psi^2}{1-\psi+\psi^2}) & -\frac{\psi(1-\psi)(\beta_2-\mu_M)}{1-\psi+\psi^2} \\ -\frac{\psi(1-\psi)(\beta_2-\mu_M)}{1-\psi+\psi^2} & (\beta_2 - \mu_M)\left(1 - \frac{2-2\psi+\psi^2}{1-\psi+\psi^2}\right) \end{bmatrix}.$$

The eigenvalues of this are given by

$$\begin{aligned} \lambda_1 =& \frac{\beta_2 - \mu_M}{2(1-\psi_1+\psi_1^2)}[-1 + \sqrt{1 - \psi_1(1-\psi_1)(1-\psi_1(1-\psi_1))}] \\ \lambda_2 =& \frac{\beta_2 - \mu_M}{2(1-\psi_1+\psi_1^2)}[-1 - \sqrt{1 - \psi_1(1-\psi_1)(1-\psi_1(1-\psi_1))}], \end{aligned}$$

which both have negative real part if $\beta_2 > \mu_M$ and $0 < \psi, (1-\psi) < 1$. As such either both genotypes will die out or both will survive depending on the values of β_2 and μ_M. The interspecific competition rates are given by $D_M(1-\psi)$ and $D_M\psi$, with $0 < \psi < 1$, while the intraspecific competition rates are both given by D_M. As the rates of intraspecific competition are greater than the interspecific rates the analysis of the Lokta-Volterra model says that the two species would coexist, which supports the results shown here.

Unlike the basic mean-field model in (3.8), this model may have equilibria that are stable yet also reactive as it has more than one variable. In order to determine if this is the case we find the

Hermitian matrix using $H = 1/2(J + J^T)$ which gives

$$H = \begin{bmatrix} \beta_2 - \mu_M - D_M(2M_S + (1-\psi)M_R) & \frac{-D_M((1-\psi)M_S + \psi M_R)}{2} \\ \frac{-D_M((1-\psi)M_S + \psi M_R)}{2} & \beta_2 - \mu_M - D_M(2M_R + \psi M_S) \end{bmatrix}.$$

Only the reactivity of the stable equilibria is of interest, so the investigation is limited to the equilibria where both genotypes survive. The Hermitian in this case is given by

$$H(M_S^*, M_R^*) = \begin{bmatrix} (\beta_2 - \mu_M)(1 - \frac{1+\psi^2}{1-\psi+\psi^2}) & -\frac{2\psi(1-\psi)(\beta_2-\mu_M)}{1-\psi+\psi^2} \\ -\frac{2\psi(1-\psi)(\beta_2-\mu_M)}{1-\psi+\psi^2} & (\beta_2 - \mu_M)\left(1 - \frac{2-2\psi+\psi^2}{1-\psi+\psi^2}\right). \end{bmatrix} \tag{5.8}$$

The eigenvalues of which are given by

$$\begin{aligned} \lambda_1 &= \frac{\beta_2 - \mu_M}{1-\psi+\psi^2}\left(3 + \sqrt{3(3+4\psi^2(1-\psi)^2)}\right) \\ \lambda_1 &= \frac{\beta_2 - \mu_M}{1-\psi+\psi^2}\left(3 - \sqrt{3(3+4\psi^2(1-\psi)^2)}\right) \end{aligned} \tag{5.9}$$

As $\sqrt{3(3+4\psi^2(1-\psi)^2)} > 3$ this will lead to two real valued eigenvalues, one of which will be positive and the other negative. This result suggests that under certain conditions transient growth may be seen after the system is perturbed away from equilibrium. The potential reactivity of this equilibrium means that when treatment is applied or removed the state may not go immediately back to the corresponding equilibrium. Whether it does or not is dependent on state following this perturbation, just as the initial conditions of a dynamics system affect the trajectory taken to equilibrium. This effect may or may not be beneficial to applying treatment. It may be beneficial if this took the form of a drop below the treated equilibrium either before reaching this equilibrium during treatment or immediately after treatment. In contrast an increase to above untreated equilibrium levels resulting from starting or stopping treatment would not be beneficial.

5.2.2 Two Genotype Variance Inclusive Model Analysis

The two genotype variance inclusive model, much like the single genotype variance inclusive model (3.24), is more difficult to analyse. However, using the single genotype variance inclusive model and the haploid mean-field model can help provide some insight. In chapter three, three conditions on the parameters were outlined that could lead to potentially different equilibrium states. These were $\beta_2 < \mu_M$, $\mu_M < \beta_2 < \mu_M + D_M$, and $\beta_2 > \mu_M + D_M$. When $\beta_2 > \mu_M + D_M$ a stable equilibrium with a positive mean and variance was expected. When $\beta_2 < \mu_M$ it was expected that the parasites would die out. When $\mu_M < \beta_2 < \mu_M + D_M$, the Jacobian suggested that the parasites would die out. However, the numerical simulations conducted in this research could not achieve this. These three conditions are of interest again in regards to this model when it is reduced to the single-species model when one dies out and, to see how they may change the expectations of coexistence in the model.

In the case that one genotype is entirely eradicated from the system and the corresponding mean and variance become zero our model reduces to the single genotype variance inclusive model. For notational purposes, we will examine the case that only the susceptible parasites survive. From the analysis of the single genotype model, we know that one equilibrium is the trivial equilibrium, with $(M_S, M_R, V_S, V_R) = (0, 0, 0, 0)$. As the trivial equilibrium is stable for the mean-field model for two genotypes when $\beta_2 < \mu_M$, it is expected that this would also be the case here as each of the ODEs is very similar to the mean-field equation with the only real differences being an extra negative term in each.

The most complex equilibria has all four state variables positive. To analyse this a similar method to the single genotype model. First the substitutions $V_S = M_S$ and $V_R = M_R$ are made into the ODEs given in equation (5.5) to give:

$$\frac{dM_S}{dt} = \beta_2 M_S - \mu_M M_S - D_M(M_S^2 + (1-\psi)M_S M_R + M_S)$$

$$\frac{dM_R}{dt} = \beta_2 M_R - \mu_M M_R - D_M(M_R^2 + \psi M_R M_S + M_R)$$

$$\begin{aligned}\frac{dV_S}{dt} &= \beta_2 M_S - \mu_M(2V_S - M_S) \\ &\quad - D_M(\frac{4V_S^2}{M_S} + 4M_S V_S - M_S^2 - 3V_S + (1-\psi)(2M_R V_S - M_R M_S)) \\ &= \beta_2 M_S - \mu_M(M_S) - D_M(M_S + 3M_S^2 + (1-\psi)M_R M_S)\end{aligned}$$

$$\begin{aligned}\frac{dV_R}{dt} &= \beta_2 M_R - \mu_M(2V_R - M_R) \\ &\quad - D_M(\frac{4V_R^2}{M_R} + 4M_R V_R - M_R^2 - 3V_R + \psi(2M_S V_R - M_S M_R)), \\ &= \beta_2 M_R - \mu_M(M_R) - D_M(M_R + 3M_R^2 + \psi M_S M_R).\end{aligned}$$

Similar to the analysis from chapter 3 this shows that $\frac{dV_S}{dt} < \frac{dM_S}{dt}$ and $\frac{dV_R}{dt} < \frac{dM_R}{dt}$, which implies that:

$$0 \leq V_S \leq M_S \text{ and } 0 \leq V_R \leq M_R.$$

With this the rates of change of the means when these variances are at either their upper or lower bounds may be examined. Setting $V_S \geq 0$ and $V_R \geq 0$ then the rates of change of the means become:

$$\begin{aligned}\frac{dM_S}{dt} &\leq \beta_2 M_S - \mu_M M_S - D_M(M_S + (1-\psi)M_S M_R) \\ \frac{dM_R}{dt} &\leq \beta_2 M_R - \mu_M M_R - D_M(M_R + \psi M_R M_S).\end{aligned}$$

Parameter set	Criteria	Parameter	Value
Set 1	$\beta > \mu_M + D_M$	μ_M	$\frac{1}{5\times 365}$
		β	1.5
		D_M	$\frac{1}{20}$
		ψ	0.9
Set 2	$\beta < \mu_M$	μ_M	2
		β	1.5
		D_M	$\frac{1}{20}$
		ψ	0.9
Set 3	$\mu_M < \beta < \mu_M + D_M$	μ_M	$1.5 - \frac{1}{30}$
		β	1.5
		D_M	$\frac{1}{20}$
		ψ	0.9

Table 5.1: Parameter sets meeting the three conditions for the variance inclusive model

By the same reasoning as the single genotype models this shows that the values of the means at equilibrium will be less than or equal to the values of the means at equilibrium given by the mean-field model. That is :

$$M_S \leq \frac{\psi(\beta_2 - \mu_M)}{D_M(1-\psi+\psi^2)} \text{ and } M_R \leq \frac{(1-\psi)(\beta_2 - \mu_M)}{D_M(1-\psi+\psi^2)}.$$

However this will only be the case if $\beta_2 > \mu_M$. Considering the upper bound on the variances, $V_S \leq M_S$ and $V_R \leq M_R$, results in:

$$\begin{aligned}
\frac{dM_S}{dt} &\geq \beta_2 M_S - \mu_M M_S - D_M M_S - D_M(M_S + (1-\psi)M_S M_R) \\
\frac{dM_R}{dt} &\geq \beta_2 M_R - \mu_M M_R - D_M M_R - D_M(M_R + \psi M_R M_S),
\end{aligned}$$

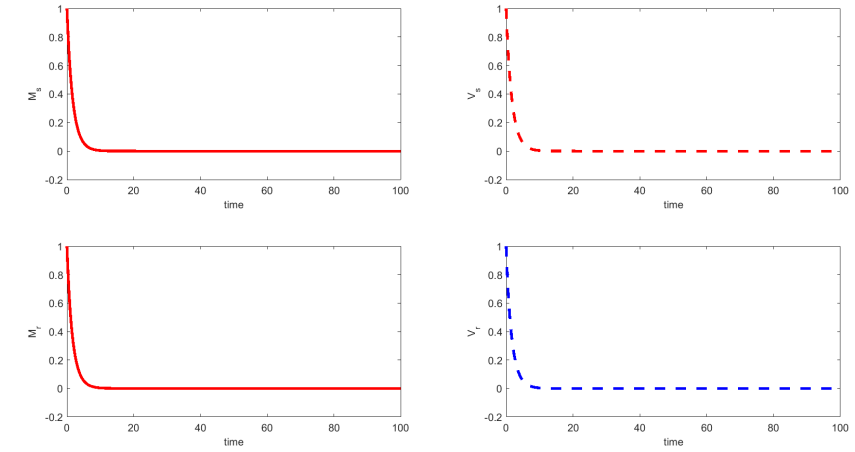
which leads to the following inequality for the equilibrium values of the means

$$M_S \geq \frac{\psi(\beta_2 - \mu_M - D_M)}{D_M(1-\psi+\psi^2)} \text{ and } M_R \geq \frac{(1-\psi)(\beta_2 - \mu_M - D_M)}{D_M(1-\psi+\psi^2)}.$$

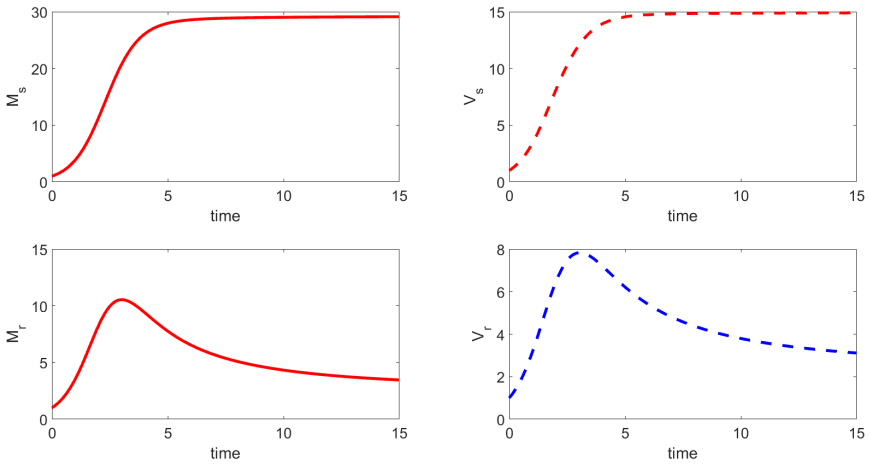
This work indicates that if $\beta_2 > \mu_M + D_M$ there will exist an equilibrium with both the means of the genotypes being positive. It is once again uncertain what will happen if $\mu_M < \beta_2 < \mu_M + D_M$, however it will be explored numerically once again.

Given the analytical difficulty involved in studying this, further numerical examples are used instead. Parameter sets which satisfy the three conditions are given in table 5.1. Figure 5-1 shows simulations of the model for each of these parameters sets. The choice of parameters is somewhat arbitrary and made simply to demonstrate behaviours. Other parameter choices would also demonstrate qualitatively similar results.

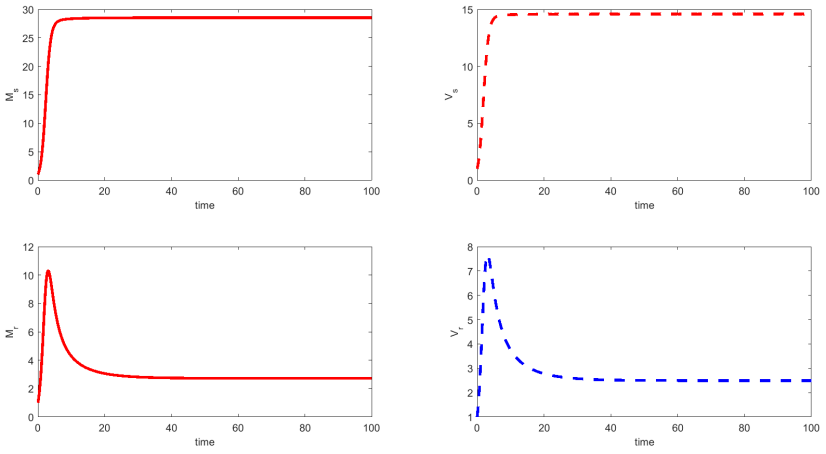
Performing these simulations showed that $\beta_2 > \mu_M + D_M$ and $\beta_2 < \mu_M$ mostly performed as



(a)



(b)



(c)

Figure 5-1: Numerical simulation to the equilibrium of the variance inclusive two genotype model with parameters given in table 5.1 for (a) Parameter set 1 (b) Parameter set 2 (c) Parameter set 3

expected tending to either coexistence or total eradication respectively. When $\mu_M < \beta_2 < \mu_M + D_M$ then the simulations also tended towards coexistence. To try and shed further light on it the stability of the equilibria are considered using the Jacobian. The Jacobian is calculated as:

$$J = \begin{bmatrix} J_{1,1} & J_{1,2} & J_{1,3} & J_{1,4} \\ J_{2,1} & J_{2,2} & J_{2,3} & J_{2,4} \\ J_{3,1} & J_{3,2} & J_{3,3} & J_{3,4} \\ J_{4,1} & J_{4,2} & J_{4,3} & J_{4,4} \end{bmatrix},$$

where the elements of the matrix are given by

$$\begin{aligned}
J_{1,1} &= \beta_2 - \mu_M - D_M(2M_S + (1-\psi)M_R) \\
J_{1,2} &= -D_M(1-\psi)MS \\
J_{1,3} &= -D_M \\
J_{1,4} &= 0 \\
J_{2,1} &= -D_M\psi M_R \\
J_{2,2} &= \beta_2 - \mu_M - D_M(2M_R + \psi M_S) \\
J_{2,3} &= 0 \\
J_{2,4} &= -D_M \\
J_{3,1} &= \beta_2 + \mu_M - D_M(\frac{-4V_S^2}{M_S^2} + 4V_S - 2M_S - (1-\psi)M_R) \\
J_{3,2} &= -D_M(1-\psi)(2V_S - M_S) \\
J_{3,3} &= -2\mu_M - D_M(\frac{8V_S}{M_S} + 4M_S - 3 + 2(1-\psi)M_R) \\
J_{3,4} &= 0 \\
J_{4,1} &= -D_M\psi(2V_R - M_R) \\
J_{4,2} &= \beta_2 + \mu_M - D_M(\frac{-4V_R^2}{M_R^2} + 4V_R - 2M_R - \psi M_S) \\
J_{4,3} &= 0 \\
J_{4,4} &= -2\mu_M - D_M(\frac{8V_R}{M_R} + 4M_R - 3 + 2\psi M_S).
\end{aligned}$$

To explore the stability of the equilibria, the parameter sets given earlier in table 5.1 are used to calculate the potential equilibria and the associated eigenvalues with the Jacobian matrix. This may be seen in table 5.2. For the parameter sets 1 and 2 with $\beta_2 > \mu_M + D_M$ and $\beta_2 < \mu_M$ respectively, the eigenvalues confirm what the simulations suggested. The equilibrium in which all four state variables are positive, and the two genotypes coexist, is stable when $\beta_2 > \mu_M + D_M$. The trivial equilibrium is stable when $\beta_2 < \mu_M$ as hypothesised based on the mean-field model. For parameter set 3, where $\mu_M < \beta_2 < \mu_M + D_M$, the trivial equilibrium was unstable, contrary to the results of the single genotype model.

Parameter set	Equilibrium (M_S, M_R, V_S, V_R)	Eigenvalues of J
Set 1	$(29.19,\ 2.79,\ 14.91,\ 2.56)$	$-5.89,\ -2.02,\ -1.46,\ -0.15$
	$(0, 0, 0, 0)$	$1.50,\ 1.50,\ 0.15,\ 0.15$
	$(29.48,\ 0,\ 15.00,\ 0)$	$-5.93,\ -1.47,\ -1.06,\ 0.06$
	$(0,\ 29.48,\ 0,\ 15.00)$	$-5.93,\ -1.47,\ 1.29,\ 0.07$
Set 2	$(0, 0, 0, 0)$	$-0.025 \pm 0.37i$
Set 3	$(0, 0, 0, 0)$	$0.15,\ 0.15,\ -0.03,\ -0.03$
	$(28.81,\ 0, 14.66,\ 0)$	$-5.80,\ -1.43,\ -1.03,\ 0.05$
	$(0,\ 28.81,\ 0,\ 14.66)$	$-5.80,\ -1.43,\ 1.26,\ 0.07$
	$(28.54,\ 2.72,\ 14.59,\ 2.50)$	$-5.76,\ -1.97,\ -1.43,\ -0.15$

Table 5.2: Eigenvalues of the Jacobian determined at each of the equilibria for the variance inclusive two genotype model for the parameter sets shown in table 5.1.

As for the previous models the reactivity of the system at each of these equilibria are properties of interest. The Hermitian in this case is given by

$$H(J) = \begin{bmatrix} H_{1,1} & H_{1,2} & H_{1,3} & H_{1,4} \\ H_{2,1} & H_{2,2} & H_{2,3} & H_{2,4} \\ H_{3,1} & H_{3,2} & H_{3,3} & H_{3,4} \\ H_{4,1} & H_{4,2} & H_{4,3} & H_{4,4} \end{bmatrix},$$

with the elements of the matrix given by

$$\begin{aligned}
H_{1,1} &= \beta_2 - \mu_M - D_M(2M_S + (1-\psi)M_R) \\
H_{1,2} &= -\frac{D_M}{2}((1-\psi)M_S + \psi M_R) \\
H_{1,3} &= \frac{\beta_2 + \mu_M - D_M - D_M(\frac{-4V_S^2}{M_S^2} + 4V_S - 2M_S - (1-\psi)M_R)}{2} \\
H_{1,4} &= \frac{-D_M\psi(2V_R - M_R)}{2} \\
H_{2,1} &= -\frac{D_M}{2}((1-\psi)M_S + \psi M_R) \\
H_{2,2} &= \beta_2 - \mu_M - D_M(2M_R + \psi M_S) \\
H_{2,3} &= \frac{-D_M(1-\psi)(2V_S - M_S)}{2} \\
H_{2,4} &= \frac{\beta_2 + \mu_M - D_M - D_M(\frac{-4V_R^2}{M_R^2} + 4V_R - 2M_R - \psi M_S)}{2} \\
H_{3,1} &= \frac{\beta_2 + \mu_M - D_M - D_M(\frac{-4V_S^2}{M_S^2} + 4V_S - 2M_S - (1-\psi)M_R)}{2} \\
H_{3,2} &= \frac{-D_M(1-\psi)(2V_S - M_S)}{2} \\
H_{3,3} &= -2\mu_M - D_M(\frac{8V_S}{M_S} + 4M_S - 3 + 2(1-\psi)M_R)
\end{aligned}$$

Parameter set	Equilibrium (M_S, M_R, V_S, V_R)	Eigenvalue sf $H(J)$
Set 1	$(29.19,\ 2.79,\ 14.91,\ 2.56)$	$-6.02,\ -2.77,\ -1.33,\ 0.60$
	$(0, 0, 0, 0)$	$1.50,\ 1.50,\ 0.15,\ 0.15$
	$(29.48,\ 0,\ 15.00,\ 0)$	$-6.06,\ -2.05,\ -1.33,\ 1.04$
	$(0,\ 29.48,\ 0,\ 15.00)$	$-3.13,\ -2.73,\ -0.02,\ -0.02$
Set 2	$(0, 0, 0, 0)$	$1.55, -1.51$
Set 3	$(0, 0, 0, 0)$	$0.15,\ 0.15,\ -0.04,\ -0.04$
	$(28.81,\ 0, 14.66,\ 0)$	$-5.93,\ -1.999,\ -1.30,\ 1.02$
	$(0,\ 28.81,\ 0,\ 14.66)$	$-5.93,\ 1.80,\ -1.48,\ -0.30$
	$(28.54,\ 2.72,\ 14.59,\ 2.50)$	$-5.90,\ -2.71,\ -1.30,\ 0.59$

Table 5.3: Eigenvalues of the Hermitian determined at each of the equilibria for the variance inclusive two genotype model for the parameter sets shown in table 5.1.

$$
\begin{aligned}
H_{3,4} &= 0 \\
H_{4,1} &= \frac{-D_M\psi(2V_R - M_R)}{2} \\
H_{4,2} &= \frac{\beta_2 + \mu_M - D_M - D_M(\frac{-4V_R^2}{M_R^2} + 4V_R - 2M_R - \psi M_S)}{2}
\end{aligned}
$$

$$
\begin{aligned}
H_{4,3} &= 0 \\
H_{4,4} &= -2\mu_M - D_M(\frac{8V_R}{M_R} + 4M_R - 3 + 2\psi M_S).
\end{aligned}
$$

Using this matrix with the different equilibrium values, that were determined numerically for the three parameter sets in table 5.1, leads to the results shown in table 5.3. What may be seen from these results is that, for the stable equilibria, the eigenvalues are a mix of positive and negative. As a result of this there is potential for transient growth of the model following a perturbation away from these equilibria, prior to returning to it, although it is not guaranteed. This is because reactivity is given by the strongest initial response to a perturbation. While under certain perturbations this initial response may lead to transient growth, for others it may not. The sign of the dominant eigenvalue may indicate whether transient growth is likely, with a positive sign suggesting transient growth will occur. When not all eigenvalues share the same sign it is possible that for some perturbations the model may not act in the expected way. If we look at the non-stable equilibria we see that for parameter set 1, where $\beta_2 > \mu_M + D_M$ all the eigenvalues are positive, which implies that the state variables will grow away from the trivial equilibrium as expected.

5.3 Diploid Model

The second model presented in this chapter is an advancement on the models where it is no longer assumed that the parasite genotype is dictated by a single allele. In this model it is assumed that each parasite has a genotype that is determined by two alleles, known as diploid genetics[83]. This model is more complex than the haploid models as it is no longer assumed that each parasite reproduces asexually. Instead parasites reproduce as a result of a direct interaction with another parasite. The

offspring produced then inherit a single allele from each parent.

5.3.1 Model Formulation

The basis of this model has some similarities to the other models. It begins with a host population of size N, indexed by i. Within the i-th host there are three different populations of parasites $[M_{i,SS}]$, $[M_{i,SR}]$ and $[M_{i,RR}]$. Any parasites born as a result of an infection within a host first pass into the corresponding larval populations L_{SS}, L_{SR} and L_{RR}. The following biological processes, assumed to be Poisson processes, dictate the change in the size of the populations.

- Birth: All parasites within the i-th host reproduce at a rate dependent on the number parasites available in the host. The rates at which the mature parasites mate and the offspring move into the correct larval category are given by $f_1([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$, $f_2([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$ and $f_3([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$ for the SS, SR and RR parasites respectively
- Larval death: All larvae living outside a host are subject to a linear death rate of μ_M as a result of natural causes
- Larval Pick-up: Larval stage parasites are each picked up at a rate $N\phi$ with the probability of each host being the one to pick them up given by $\frac{1}{N}$.
- Larval Maturation: Once in the host the probability that larvae mature is given by ρ with those that don't dying. The maturation process is immediate.
- Natural parasite death: All parasites within a host are subject to a linear death rate due to natural causes of μ_M.
- Density dependent mortality: All parasites suffer a death rate due to competition in the host. The death rates are given by $g_1([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$, $g_2([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$ and $g_3([M_{i,SS}], [M_{i,SR}], [M_{i,RR}])$ for the SS, SR and RR parasites respectively.

The density dependent parasite death rate is very similar to the two genotype model. For every parasite in a host all parasites will suffer an additional death rate. This is calculated as a weighted average, by multiplying the base increase by the number of each genotype parasites present and multiplying by the probability that one genotype will survive a competition with another. We set D_M to denote the average rate at which any parasite dies due to competition with another parasite. Table 5.4 shows the probability of each genotype parasite dying in a competition with a parasite of each other genotype and the resulting death rate that those parasites inflict on one another. Summing these death rates over all the parasites within the i-th host gives the total increased death rates for each parasite of a set genotype. These total death rates on a single parasite are as follows:

- On an SS parasite in the i-th host: $D_M([M_{i,SS}] + (1-\psi_1)[M_{i,SR}] + (1-\psi_2)[M_{i,RR}])$.
- On an SR parasite in the i-th host: $D_M([M_{i,SR}] + \psi_1[M_{i,SS}] + (1-\psi_3)[M_{i,RR}])$.

Genotypes Competing	Loser	Probability of death	Increased death rate per competitor
$SS-SS$	SS	1	D_M
$SR-SR$	SR	1	D_M
$RR-RR$	RR	1	D_M
$SS-SR$	SR	ψ_1	$\psi_1 D_M$
$SS-SR$	SS	$1-\psi_1$	$(1-\psi_1)D_M$
$RR-SR$	SR	$1-\psi_3$	$(1-\psi_3)D_M$
$RR-SR$	RR	ψ_3	$\psi_3 D_M$
$SS-RR$	SS	$1-\psi_2$	$(1-\psi_2)D_M$
$SS-RR$	RR	ψ_2	$\psi_2 D_M$

Table 5.4: Probability of each genotype parasite dying in competition with a parasite of another genotype and the increase that the each of surviving genotype parasites inflicts on the non-surviving genotype parasites.

- On an RR parasite in the i-th host: $D_M([M_{i,RR}] + \psi_3[M_{i,SR}] + \psi_2[M_{i,SS}])$.

As with the haploid model, we take a very simplistic approach here to establish a coupled system which will be used to explore the impact of control when there is population variability within a host. The birth rate in this model is more complex than any of the other models as we assume that parasites may reproduce as a result of self fertilisation or via sexual reproduction, with an equal probability of a parasite mating with any other parasite or self fertilising. First we consider the probability of a parasite reproducing as a result of an interaction with each of the other genotypes in a host with. For any of the parasites which either mate with another parasite or self fertilise in order to lay eggs the probability that half of any larval offspring's genetics is provided by each of the different genotypes as:

- From SS parasite: $\frac{[M_{i,SS}]}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$
- From SR parasite: $\frac{[M_{i,SR}]}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$
- From RR parasite: $\frac{[M_{i,RR}]}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$

where $[M_{i,SS}]$, $[M_{i,SR}]$ and $[M_{i,RR}]$ denote the burden within the host.

However this is only part of the overall birth rate. We must also consider the probability that the resulting offspring of each mating or self fertilisation are born into each of the three larval populations is also important. Assuming that each allele is equally likely to be selected for the offspring, table 5.5 shows the probability that the offspring will belong to each genotype given the genotypes of it's parent(s). Combining the rates at which these matings occur and the probability of the offspring born being each genotype gives that the overall birth rate of each genotype of larval parasite from the parasites within the i-th host are given by

- Birth rate of SS: $\beta\frac{([M_{i,SS}^2]+[M_{i,SR}][M_{i,SS}]+0.25[M_{i,SR}]^2)}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$.
- Birth rate of SR: $\beta\frac{([M_{i,SR}][M_{i,SS}]+0.5[M_{i,SR}]^2+[M_{i,SR}][M_{i,RR}])+2[M_{i,SS}][M_{i,RR}]}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$.

Pairing	Probability of SS larva	Probability of SR larva	Probability of RR larva
$SS-SS$	1	0	0
$SS-SR$	0.5	0.5	0
$SS-RR$	0	1	0
$SR-RR$	0	0.5	0.5
$SR-SR$	0.25	0.5	0.25
$RR-RR$	0	0	1

Table 5.5: Probability of different genotype larval offspring being born as a result of each possible pairing of mature parasites reproducing.

- Birth rate of RR: $\beta\frac{([M_{i,RR}^2]+[M_{i,SR}][M_{i,RR}]+0.25[M_{i,SR}]^2)}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$.

With these processes better defined the model formulation may be begun, by following the same methodology as before. Doing this gives the initial ODEs for the rate of change of the sizes of each population within the i-th host and larval populations as:

$$\begin{aligned}
\frac{dL_{SS}}{dt} &= \beta\sum_{j=1}^{N}\frac{([M_{j,SS}]^2+[M_{j,SR}][M_{j,SS}]+0.25[M_{j,SR}]^2)}{[M_{j,SS}]+[M_{j,SR}]+[M_{j,RR}]}-\mu_L L_{SS}-N\phi L_{SS}\\
\frac{dL_{SR}}{dt} &= \beta\sum_{j=1}^{N}\frac{([M_{j,SR}][M_{j,SS}]+0.5[M_{j,SR}]^2+[M_{j,SR}][M_{j,RR}])+2[M_{j,SS}][M_{j,RR}]}{[M_{j,SS}]+[M_{j,SR}]+[M_{j,RR}]}\\
&\quad -\mu_L L_{SR}-N\phi L_{SR}\\
\frac{dL_{RR}}{dt} &= \beta\sum_{j=1}^{N}\frac{([M_{j,RR}]^2+[M_{j,SR}][M_{j,RR}]+0.25[M_{j,SR}]^2)}{[M_{j,SS}]+[M_{j,SR}]+[M_{j,RR}]}-\mu_L L_{RR}-N\phi L_{RR}\\
\frac{d[M_{i,SS}]}{dt} &= \rho\phi L_{SS}-\mu_M[M_{i,SS}]-D_M[M_{i,SS}]([M_{i,SS}]+(1-\psi_1)[M_{i,SR}]+(1-\psi_2)[M_{i,RR}])\\
\frac{d[M_{i,SR}]}{dt} &= \rho\phi L_{SR}-\mu_M[M_{i,SR}]-D_M[M_{i,SR}]([M_{i,SR}]+\psi_1[M_{i,SS}]+(1-\psi_3)[M_{i,RR}])\\
\frac{d[M_{i,RR}]}{dt} &= \rho\phi L_{RR}-\mu_M[M_{i,RR}]-D_M[M_{i,RR}]([M_{i,RR}]+\psi_3[M_{i,SR}]+\psi_2[M_{i,SS}]).
\end{aligned}$$

Previously the birth rate of the larvae simplified at this point to a function of the mean. This does not apply here as it is no longer a linear function or even simply a polynomial function. The birth rate now involves the reciprocal of the sum of the state variables, that is $\frac{1}{[M_{i,SS}]+[M_{i,SR}]+[M_{i,RR}]}$. In general for randomly distributed $\{X_i\}$,

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{X_i}\neq\frac{1}{\frac{1}{N}\sum_{j=1}^{N}X_i}.$$

In fact it may even be that $\frac{1}{N}\sum_{j=1}^{N}\frac{1}{X_i}$ is undefined if any X_i takes the value zero. As the variables $[M_{i,SS}],\ [M_{i,SR}]$ and $[M_{i,RR}]$ may take the value zero this causes a problem in the model. The way that this is overcome is to limit the scope of this model to only a mean-field model such that it may be assumed that $[M_{i,SS}]=\overline{M}_{SS},\ [M_{i,SR}]=\overline{M}_{SR}$ and $[M_{i,RR}]=\overline{M}_{RR}$ for all i. Using this, and

averaging across all simulations, as in the methods discussed in Chapter 3, the model becomes

$$\begin{aligned}
\frac{dL_{SS}}{dt} &= \beta N \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - \mu_L L_{SS} - N\phi L_{SS} \\
\frac{dL_{SR}}{dt} &= \beta N \frac{2M_{SS}M_{RR} + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \mu_L L_{SR} - N\phi L_{SR} \\
\frac{dL_{RR}}{dt} &= \beta N \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - \mu_L L_{RR} - N\phi L_{RR} \\
\frac{dM_{SS}}{dt} &= \rho\phi L_{SS} - \mu_M M_{SS} - D_M M_{SS}(M_{SS} + (1-\psi_1)M_{SR} + (1-\psi_2)M_{RR}) \\
\frac{dM_{SR}}{dt} &= \rho\phi L_{SR} - \mu_M M_{SR} - D_M M_{SR}(M_{SR} + (\psi_1)M_{SR} + (1-\psi_3)M_{RR}) \\
\frac{dM_{RR}}{dt} &= \rho\phi L_{RR} - \mu_M M_{RR} - D_M M_{RR}(M_{RR} + (\psi_2)M_{SS} + (\psi_3)M_{SR})
\end{aligned} \tag{5.10}$$

where $M_{SS} = \frac{1}{W}\sum_{k=1}^{W}\frac{1}{N}\sum_{i=1}^{N} M_{i,SS}$ and similar for the other two genotypes. It should also be noted that although the denominator in the birth rate $M_{SS} + M_{SR} + M_{RR}$, would suggest that the model is only valid if $M_{SS} + M_{SR} + M_{RR} > 0$. However as this tends to zero, the numerators of all birth rates will also tend to zero at a rate that is either faster or equal to the denominator. Resultantly, as the sum of all the means tends to zero so does the birth rate.

In a similar manner to the previous models this model is simplified by assuming that the rates of larval processes is typically far quicker than the mature parasites. As a result the larval populations reach equilibrium very quickly and the rates of change of the larval populations are assumed to be fixed at zero. This assumption allows the larval populations to be set such that they are equal to

$$\begin{aligned}
L_{SS} &= \frac{\beta N}{\mu_L - N\phi}\frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} \\
L_{SR} &= \frac{\beta N}{\mu_L - N\phi}\frac{2M_{SS}M_{RR} + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}M_{SR}}{M_{SS} + M_{SR} + M_{RR}} \\
L_{RR} &= \frac{\beta N}{\mu_L - N\phi}\frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}}.
\end{aligned}$$

Substituting these into the ODEs for the mean and setting $\beta_2 = \frac{N\beta\rho\phi}{\mu_L + N\phi}$ the model is reduced to

$$\begin{aligned}
\frac{dM_{SS}}{dt} &= \beta_2 \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - \mu_M M_{SS} \\
&\quad - D_M M_{SS}(M_{SS} + (1-\psi_1)M_{SR} + (1-\psi_2)M_{RR}) \\
\frac{dM_{SR}}{dt} &= \beta_2 \frac{2M_{SS}M_{RR} + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \mu_M M_{SR} \\
&\quad - D_M M_{SR}(M_{SR} + (\psi_1)M_{SR} + (1-\psi_3)M_{RR}) \\
\frac{dM_{RR}}{dt} &= \beta_2 \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - \mu_M M_{RR} \\
&\quad - D_M M_{RR}(M_{RR} + (\psi_2)M_{SS} + (\psi_3)M_{SR}).
\end{aligned} \tag{5.11}$$

Although the model is still not tractable this is done for similarity with the other models so that comparisons may be better demonstrated.

5.3.2 Numerical Simulation

Aside from the trivial equilibria, the easiest equilibria to establish occurs when only one of the homozygous genotypes remains in the model, with the heterozygous genotype and the other homozygous genotype both eliminated. When this occurs the model is reduced to the basic model as neither of the other genotype parasites may be born as a result of the infection by the single homozygous genotype present. As a result of this it is known that equilibria will occur when

$$(M^*_{SS}, M^*_{SR}, M^*_{RR}) = (\frac{\beta_2 - \mu_M}{D_M}, 0, 0) \text{ and } (M^*_{SS}, M^*_{SR}, M^*_{RR}) = (0, 0, \frac{\beta_2 - \mu_M}{D_M}).$$

In addition when looking at the equilibria in this way we can see that there will also be an equilibrium at

$$(M^*_{SS}, M^*_{SR}, M^*_{RR}) = (0, 0, 0),$$

due to the birth rate tending to zero as the sum of the means tends to zero. If $\beta_2 > \mu_M$ then we know from the analysis of the basic model that we will have two non-zero equilibria. Consider now a state where both the homozygous genotypes are eliminated, and $M_{SS} = M_{RR} = 0$, but the heterozygous genotype M_{SR} is still present. The ODEs become

$$\begin{aligned}
\frac{dM_{SS}}{dt} &= \beta_2 0.25 M_{SR} \\
\frac{dM_{SR}}{dt} &= \beta_2 0.5 M_{SR} - \mu_M M_{SR} - D_M M_{SR}(M_{SR}) \\
\frac{dM_{RR}}{dt} &= \beta_2 0.25 M_{SR}.
\end{aligned}$$

Which shows that the heterozygous parasites cannot be the only parasites present in the population.

The other possibility is that we have an equilibrium where all three genotypes coexist. Due to the complexity of the model we cannot find an analytic representation of what this, or these, may be so instead we look at simulations for different parameter sets to determine if this is a stable equilibria that arises. The expectation based on the analysis of the system when any two genotypes remain is that this equilibrium will not just exist but be stable.

From our analysis of potential single genotype equilibria we have identified two conditions of interest on our parameter sets:

- $\mu_M > \beta_2$

- $\beta_2 > \mu_M$

Parameter set	Condition	Parameter	Value
Set 1	$\mu_M > \beta_2$	β_2	1.5
		μ_M	2
		D_M	$\frac{1}{20}$
		ψ_1	0.7
		ψ_2	0.9
		ψ_3	0.6
Set 2	$\mu_M < \beta_2$	β_2	1.5
		μ_M	$\frac{1}{5\times 365}$
		D_M	$\frac{1}{20}$
		ψ_1	0.7
		ψ_2	0.9
		ψ_3	0.6

Table 5.6: Parameter sets for the Diploid mean-field model which satisfy the conditions of interest identified during steady state analysis.

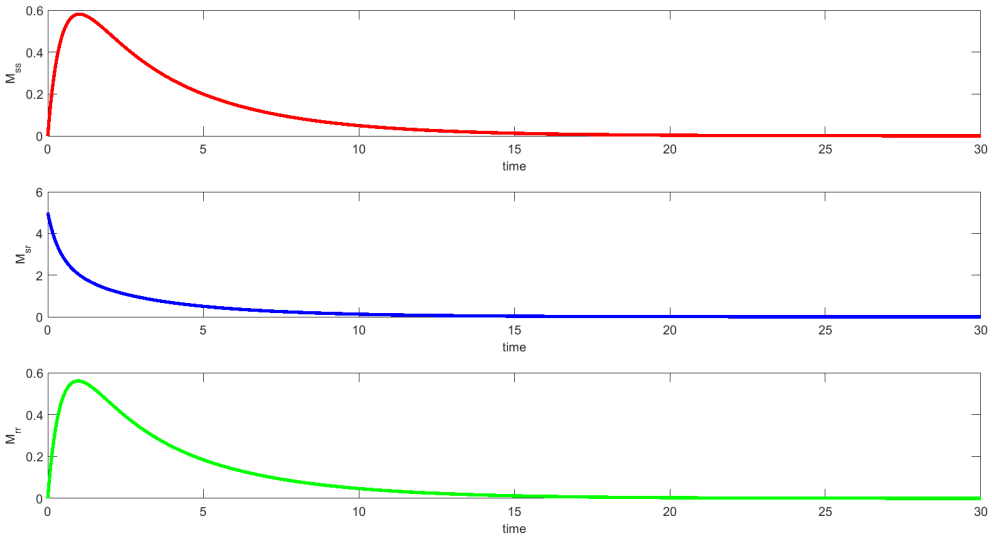
Table 5.6 gives three parameter sets which each satisfy one of these conditions. Using these parameter sets we simulate the model dynamics, which is shown in figure 5-2. These figures show that as expected when $\mu_M > \beta_2$ the model tends to the trivial equilibrium. As sub-figure (b) shows when $\beta_2 > \mu_M$ the model tends to an equilibrium in which all three genotypes coexist at varying levels. These simulations suggest that unless the initial conditions of the system are such that only one of the homozygous genotype populations is non zero and both other populations are zero then the model will not tend to an equilibria with only a single genotype population remaining.

To analyse the stability we determine the Jacobian of the system as

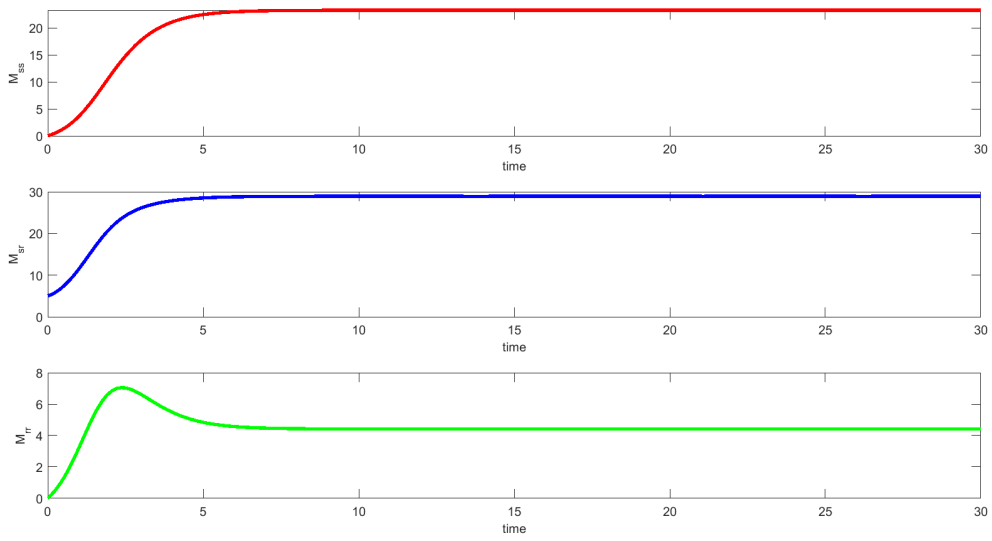
$$J = \begin{bmatrix} J_{1,1} & J_{1,2} & J_{1,3} \\ J_{2,1} & J_{2,2} & J_{2,3} \\ J_{3,1} & J_{3,2} & J_{3,3} \end{bmatrix},$$

where

$$\begin{aligned}
J_{1,1} &= \beta_2 \left(\frac{2M_{SS} + M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
&\quad - D_M(2M_{SS} + (1-\psi_1)M_{SR} + (1-\psi_2)M_{RR}) \\
J_{1,2} &= \beta_2 \left(\frac{M_{SS} + 0.5M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - D_M(1-\psi_1)M_{SS} \\
J_{1,3} &= \beta_2 \left(-\frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - D_M(1-\psi_2)M_{SS} \\
J_{2,1} &= \beta_2 \left(\frac{M_{SR} + 2M_{RR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{2M_{SS}M_{RR} + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - D_M\psi_1 M_{SR}
\end{aligned}$$



(a)



(b)

Figure 5-2: Model dynamics tending to equilibrium for parameters given in table 5.6 for (a) Parameter set 1 (b) Parameter set 2

Parameter set	Equilibrium $(M^*_{SS}, M^*_{SR}, M^*_{RR})$	Eigenvalues of J
Set 2	$(23.29, 28.93, 4.42)$	$-2.65 + 0.52i,\ -2.65 - 0.52i,\ -1.1 - +0.00i$
	$(29.99, 0, 0)$	$-2.85,\ -1.50,\ 0.45$
	$(0, 0, 29.99)$	$0.90,\ -0.15,\ 0.001$

Table 5.7: Equilibrium states and the corresponding eigenvalues of the Jacobian for the diploid mean-field model for the parameter sets given in table 5.6.

$$
\begin{aligned}
J_{2,2} &= \beta_2 \left(\frac{M_{RR} + M_{SS} + M_{SR}}{M_{RR} + M_{SS} + M_{SR}} - \frac{2M_{SS}M_{RR} + M_{SR}(M_{SS} + M_{RR} + 0.5M_{SR})}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
&\quad - D_M(2M_{SR} + \psi_1 M_{SS} + (1 - \psi_3)M_{RR}) \\
J_{2,3} &= \beta_2 \left(\frac{M_{SR} + 2M_{SS}}{M_{SS} + M_{SR} + M_{RR}} - \frac{2M_{SS}M_{RR} + M_{SR}(M_{SS} + M_{RR} + 0.5M_{SR})}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - D_M(1 - \psi_3)M_{SR} \\
J_{3,1} &= \beta_2 \left(-\frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - D_M \psi_2 M_{RR} \\
J_{3,2} &= \beta_2 \left(\frac{M_{RR} + 0.5M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - D_M \psi_3 M_{RR} \\
J_{3,3} &= \beta_2 \left(\frac{2M_{RR} + M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
&\quad - D_M(2M_{RR} + \psi_3 M_{SR} + \psi_2 M_{SS}).
\end{aligned}
$$

Using the same parameter sets we determined the Jacobian and its eigenvalues numerically to give the stability of the different equilibria that could be found numerically. The outcome of this is shown in table 5.7, it should be noted that due to the nature of the Jacobian results are not shown for the trivial equilibrium as the resulting matrix will either contain elements with denominators of zero or will be simplified down to a matrix of zeros. This would show that the equilibrium is non-hyperbolic and it's stability cannot be determined from the sign of the eigenvalues. What the results of this show is that for both parameter set two, shown in table 5.6, the equilibrium in which all three genotypes remain is stable while the equilibria with only a single homozygous population surviving are unstable. This confirms that, unless the initial conditions allow for it, no singular genotype of parasites will be entirely eradicated.

Finally in our analysis of the model we use the Hermitian $H = \frac{1}{2}(J + J^T)$ and the eigenvalues of this evaluated at the equilibria to consider any potential reactivity of the model. Our Hermitian for this model is given by

$$
H = \begin{bmatrix} H_{1,1} & H_{1,2} & H_{1,3} \\ H_{2,1} & H_{2,2} & H2,3 \\ H_{3,1} & H_{3,2} & H_{3,3} \end{bmatrix},
$$

Parameter set	Equilibrium $(M^*_{SS}, M^*_{SR}, M^*_{RR})$	Eigenvalues of H
Set 2	$23.2928.934.42$	$-3.24,\ -2.16,\ -1.01$
	$(29.99, 0, 0)$	$-3.24,\ -1.13,\ 0.48$
	$(0, 0, 29.99)$	$2.003,\ -1.53,\ 0.27$

Table 5.8: Equilibrium states and the corresponding eigenvalues of the Hermitian for the diploid mean-field model for some of the parameter sets given in table 5.6.

where

$$
\begin{aligned}
H_{1,1} &= \beta_2 \left(\frac{2M_{SS} + M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
&\quad - D_M(2M_{SS} + (1-\psi_1)M_{SR} + (1-\psi_2)M_{RR}) \\
H_{1,2} &= \frac{\beta_2}{2} \left(\frac{M_{SS} + 1.5M_{SR} + 2M_{RR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{2M_{SS}M_{RR} + M_{SS}^2 + 2M_{SS}M_{SR} + 0.75M_{SR}^2 + M_{RR}M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - \frac{D_M}{2}((1-\psi_1)M_{SS} + \psi_1 M_{SR}) \\
H_{1,3} &= \frac{\beta_2}{2} \left(-\frac{M_{SS}^2 + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}^2 + M_{RR}M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \frac{D_M}{2}((1-\psi_2)M_{SS} + \psi_2 M_{RR}) \\
H_{2,1} &= \frac{\beta_2}{2} \left(\frac{M_{SS} + 1.5M_{SR} + 2M_{RR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{2M_{SS}M_{RR} + M_{SS}^2 + 2M_{SS}M_{SR} + 0.75M_{SR}^2 + M_{RR}M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - \frac{D_M}{2}((1-\psi_1)M_{SS} + \psi_1 M_{SR}) \\
H_{2,2} &= \beta_2 \left(\frac{M_{RR} + M_{SS} + M_{SR}}{M_{RR} + M_{SS} + M_{SR}} - \frac{2M_{SS}M_{RR} + M_{SR}(M_{SS} + M_{RR} + 0.5M_{SR})}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
H_{2,3} &= \frac{\beta_2}{2} \left(\frac{1.5M_{SR} + 2M_{SS} + M_{RR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{RR}^2 + 2M_{SS}M_{RR} + M_{SR}(M_{SS} + 2M_{RR} + 0.75M_{SR})}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - \frac{D_M}{2}((1-\psi_3)M_{SR} + \psi_3 M_{RR}) \\
H_{3,1} &= \frac{\beta_2}{2} \left(-\frac{M_{SS}^2 + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}^2 + M_{RR}M_{SR}}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \frac{D_M}{2}((1-\psi_2)M_{SS} + \psi_2 M_{RR}) \\
H_{3,2} &= \frac{\beta_2}{2} \left(\frac{1.5M_{SR} + 2M_{SS} + M_{RR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{RR}^2 + 2M_{SS}M_{RR} + M_{SR}(M_{SS} + 2M_{RR} + 0.75M_{SR})}{(M_{SS} + M_{SR} + M_{RR})^2} \right) \\
&\quad - \frac{D_M}{2}((1-\psi_3)M_{SR} + \psi_3 M_{RR}) \\
H_{3,3} &= \beta_2 \left(\frac{2M_{RR} + M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{(M_{SS} + M_{SR} + M_{RR})^2} \right) - \mu_M \\
&\quad - D_M(2M_{RR} + \psi_3 M_{SR} + \psi_2 M_{SS}).
\end{aligned}
$$

Using this matrix and determining its eigenvalues for each of the equilibria shown in able 5.7 gives the corresponding eigenvalues for the matrix H shown in table 5.8. For the two parameters sets shown the stable equilibria, that is the one with all three genotypes coexisting, leads to eigenvalues of H which are all negative. This suggests that transient growth is unlikely to occur in the model for small perturbations away from this equilibria, as we discussed in chapter 3.

5.4 Discussion

In this chapter, we have covered the formulation of three different models and their analysis. The most crucial elements that we have covered here are:

- The formulation of the models.
- We have shown that when the birth rate is great enough, that is either $\beta_2 > \mu_M$ the parameters may be chosen such that there will be coexistence of the different genotypes.
- We also found that transient growth, related to the reactivity, may occur on both of the haploid models, although it was not guaranteed but would be unexpected on the diploid model for small perturbations away from the stable equilibria.

We use this basic structure in the following chapter where we find that the existence of distinct genotypes leads to different outcomes for the control of parasite populations.

Chapter 6

Control on Resistant Models

This chapter takes the models that we formulated in Chapter 5 and uses the dynamics as a basis for control models. These control models are then be used in the determination of different control strategies using some of the methods we discussed in Chapter 4 adapted to consider the resistance.

6.1 Model Dynamics Under Treatment

Before we can apply treatment to the model we must consider how the underlying model dynamics will be affected by treatment. In the single genotype models, where resistance was not included, the control was assumed to be applied to all hosts in a population and caused an increase in the mature parasite death rate when applied. We make the same assumption here, namely that treatment is applied to all hosts. However as we now have some parasites which are resistant to treatment we must give consideration to how the mature parasites are affected by the treatment.

6.1.1 Haploid Models

We assume that the susceptible parasites are fully affected by the treatment and as such their linear death rate increases by a factor u. The resistant parasites may be modelled as either fully resistant, in which case they suffer no additional death rate, or they may be simply less affected, with an increase in the death rate of size γu, where $0 < \gamma < 1$. In either case we write the model dynamics under treatment as:

$$\begin{aligned} \frac{dM_S}{dt} &= \beta_2 M_S - (\mu_M + u)M_S - D_M(M_S^2 + (1-\psi)M_S M_R) \\ \frac{dM_R}{dt} &= \beta_2 M_R - (\mu_M + \gamma u)M_R - D_M(M_R^2 + \psi M_S M_R), \end{aligned} \tag{6.1}$$

Parameter	Value
μ_M	$\frac{1}{5\times 365}$
β	1.5
D_M	$\frac{1}{20}$
ψ	0.9
u_{max}	1
γ	0

Table 6.1: Parameter set used in control examples for the Haploid models, (6.1) and (6.2)

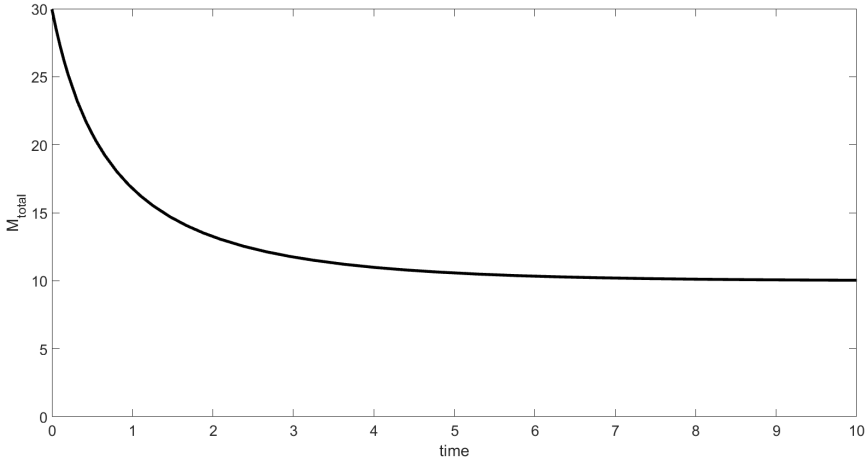
for the haploid mean-field model. For the variance inclusive haploid model, the dynamics under the influence of the control at intensity u are:

$$\begin{aligned}
\frac{dM_S}{dt} &= \beta_2 M_S - (\mu_M + u)M_S - D_M(M_S^2 + (1-\psi)M_S M_R + V_S) \\
\frac{dM_R}{dt} &= \beta_2 M_R - (\mu_M + \gamma u)M_R \\
&\quad - D_M(M_R^2 + \psi M_R M_S + V_r) \\
\frac{dV_S}{dt} &= \beta_2 M_S - (\mu_M + u)(2V_S - M_S) \\
&\quad - D_M\left(\frac{4V_S^2}{M_S} + 4M_S V_S - M_S^2 - 3V_S + (1-\psi)(2M_R V_S - M_R M_S)\right) \\
\frac{dV_R}{dt} &= \beta_2 M_R - (\mu_M + \gamma u)(2V_R - M_R) \\
&\quad - D_M\left(\frac{4V_R^2}{M_R} + 4M_R V_R - M_R^2 - 3V_R + \psi(2M_S V_R - M_S M_R)\right).
\end{aligned} \tag{6.2}$$

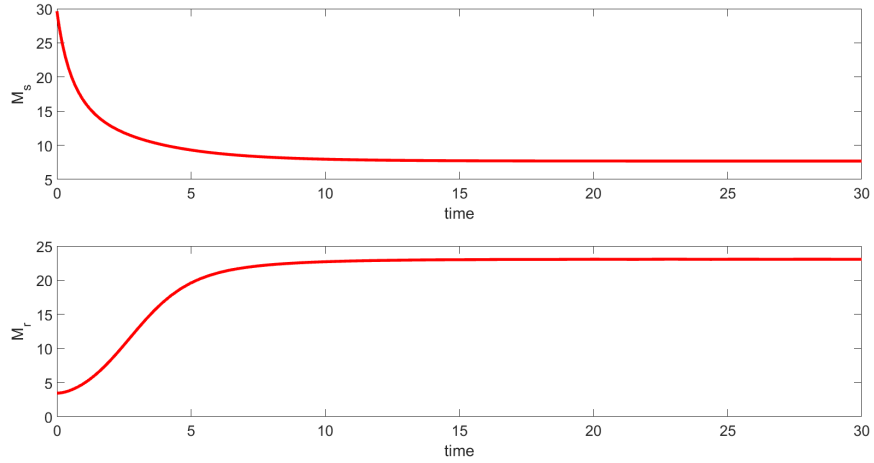
Our simulations are performed under the assumption that resistant parasites are entirely resistant, i.e with $\gamma = 0$; table 6.1 lists parameter values used throughout the chapter .

Figure 6-1 shows what occurs when a full intensity treatment is applied to the basic mean-field model and the haploid mean-field model. Comparing these, we can see that the mean parasite burden in the basic mean-field model decreases monotonically towards a lower state, the equilibrium under treatment. In contrast, when a sustained treatment is applied to the haploid mean-field model and resistance is a factor the total mean initially drops as the susceptible parasites are rapidly killed off. Following this the lack of competition from the susceptible parasites allows the resistant parasite population to build up which eventually leads to the total mean increasing again to an alternative treated equilibrium. While still lower than the untreated equilibrium this is clearly a far less favourable result.

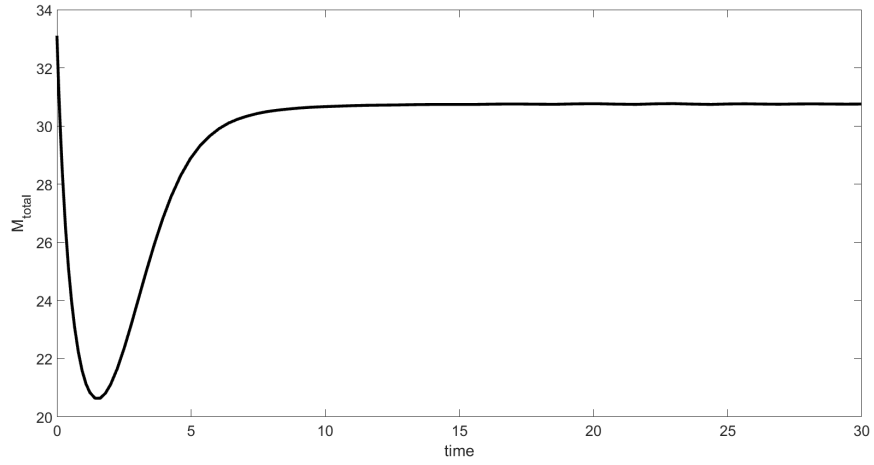
Similarly figures 6-2 and 6-3 show the application of a sustained full intensity treatment on the single genotype variance inclusive model and the two genotype variance inclusive model. We see a very similar response in these models to sustained treatment, namely the mean and variance of the parasite burden in the single genotype model monotonically decreasing to the lower treated equilibrium, and the total mean and variance of the haploid model initially decreasing before increasing.



(a)



(b)



(c)

Figure 6-1: System dynamics under full intensity treatment of (a) the basic mean-field model (4.4), for a period of 10 units for parameter values given in table 4.1 (b) and (c) the Haploid mean-field model (6.1), for a period of 30 units for parameter values given in table 6.1

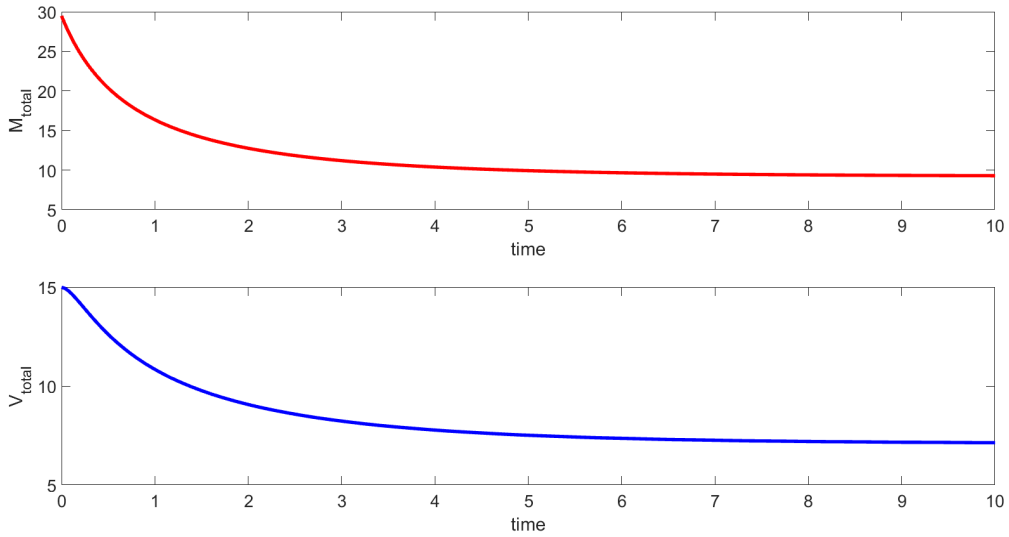


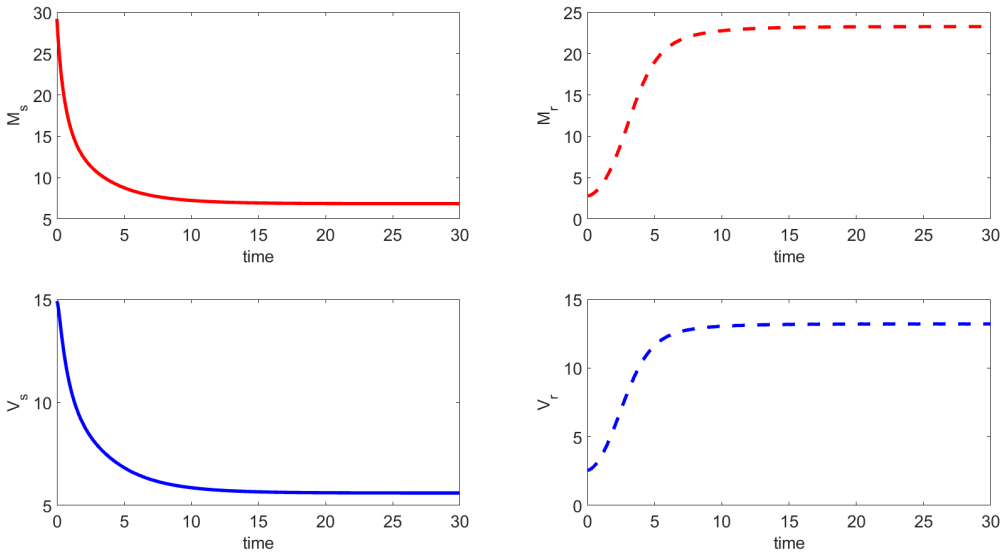
Figure 6-2: Full intensity treatment applied to the basic variance inclusive model (4.9) for a period of 10 units for parameter values given in table 4.1.

What is interesting in this case is that the Haploid variance inclusive model has a greater variance at its treated equilibrium than when not undergoing treatment. This result has implications for any treatment which is measured by its effect on the variance, as a treatment that is sustained for too long will lead to a greater overall variance than if nothing had been done.

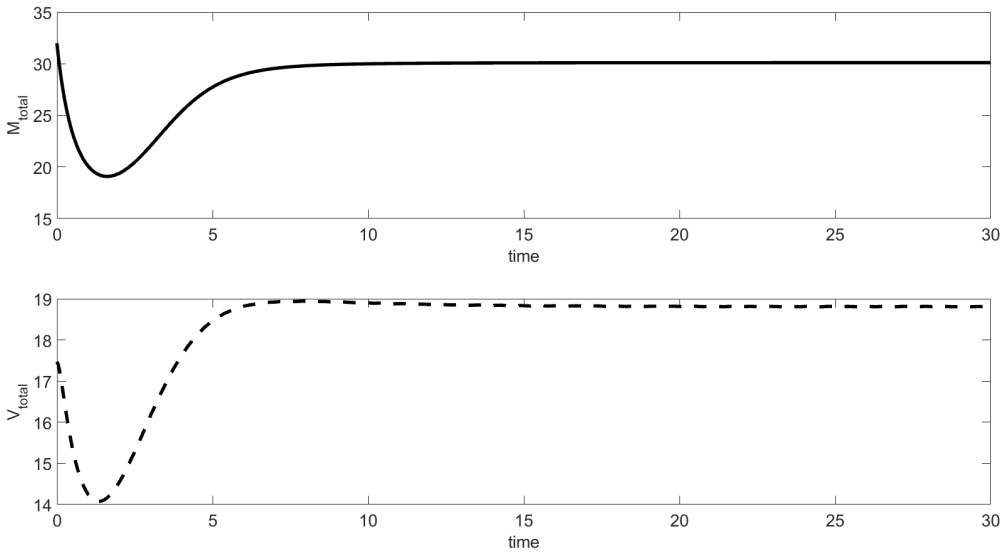
Unlike in the models which do not account for resistance, these results show that the minimal value of the total mean and variance that may be achieved does not occur at the treated equilibrium. Another interesting observation is that even if the model is not reactive on a component level, the summation of the variables $M_S + M_R$ and $V_S + V_R$ may still show transient growth. This can be seen in figures 6-1 and 6-3, sub-figures (b), for both the haploid mean-field and variance inclusive models. The effect of this property on the controls is seen throughout this chapter and the following.

6.1.2 Diploid Model

The diploid model, while similar, requires an extra level of consideration. As we now have two alleles, and as such three genotypes, we must consider how the heterozygous parasites will react to treatment. In diploid genetics, it may be the case that one gene is dominant and the presence of it will decide the phenotype of the parasites [83]. That is those which have the dominant gene, whether heterozygous or homozygous, will display the qualities associated with that gene. Alternatively, it may be that neither gene is dominant and the traits shown are due to more complex factors such as epigenetics, which concern changes in the way genes are expressed rather than the actual genetics themselves [91]. These changes may be the result of exposure to certain environmental factors and despite not changing the genes themselves may still be inherited by offspring. We make the



(a)



(b)

Figure 6-3: Full intensity treatment applied to the Haploid variance inclusive model (6.2) for a period of 10 units for parameter values given in table 6.1

.

Parameter	Value
β_2	1.5
μ_M	$\frac{1}{5\times 365}$
D_M	$\frac{1}{20}$
ψ_1	0.7
ψ_2	0.9
ψ_3	0.6
u_{max}	1
γ_1	0
γ_2	$\in \{0,1\}$ (specified for individual controls)

Table 6.2: Parameter values for the dynamics of the diploid model under treatment. The value of γ_2 is specified for each control to show the difference when heterozygous parasites act as either susceptible or resistant phenotypes.

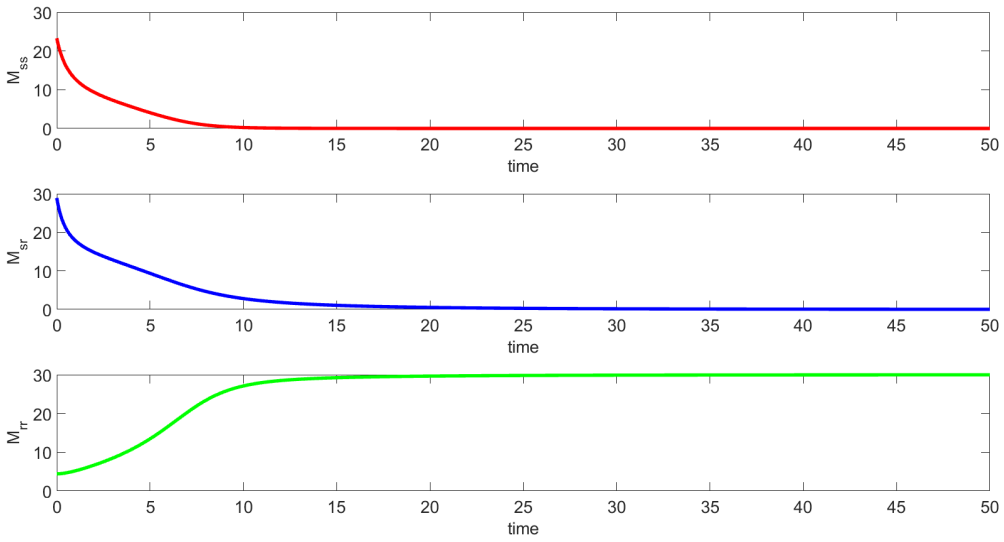
simplifying assumption that the heterozygous parasites, as a group, behave either as susceptible parasites, resistant parasites or somewhere between the two. To account for all possibilities we set the effect of treatment on susceptible parasites to be an increase in death rate of size u, the effect on resistant parasites to be an increase of size $\gamma_1 u$, and the effect on the mixed parasites to be an increase of size $\gamma_2 u$. Here $0 \leq \gamma_1 < 1$ and $\gamma_1 \leq \gamma_2 \leq 1$. It may be thought that the ability to resist treatment should be more directly related to the competitive ability, in this model the density dependent mortality rate, but as the precise nature of the resistive ability is not defined this is not considered here beyond the assumption that a genetic trade-off between increased competitive ability and resistance is made. As a result of these assumptions, the model dynamics under treatment may be written as

$$\begin{aligned}
\frac{dM_{SS}}{dt} &= \beta_2 \frac{M_{SS}^2 + M_{SS}M_{SR} + 0.25M_{SR}^2}{M_{SS}+M_{SR}+M_{RR}} - (\mu_M + u)M_{SS} \\
&\quad - D_M M_{SS}(M_{SS} + (1-\psi_1)M_{SR} + (1-\psi_2)M_{RR}) \\
\frac{dM_{SR}}{dt} &= \beta_2 \frac{2M_{SS}M_{RR} + M_{SS}M_{SR} + 0.5M_{SR}^2 + M_{RR}M_{SR}}{M_{SS}+M_{SR}+M_{RR}} - (\mu_M + \gamma_2 u)M_{SR} \\
&\quad - D_M M_{SR}(M_{SR} + (\psi_1)M_{SR} + (1-\psi_3)M_{RR}) \\
\frac{dM_{RR}}{dt} &= \beta_2 \frac{M_{RR}^2 + M_{RR}M_{SR} + 0.25M_{SR}^2}{M_{SS}+M_{SR}+M_{RR}} - (\mu_M + \gamma_1 u)M_{RR} \\
&\quad - D_M M_{RR}(M_{RR} + (\psi_2)M_{SS} + (\psi_3)M_{SR}).
\end{aligned} \tag{6.3}$$

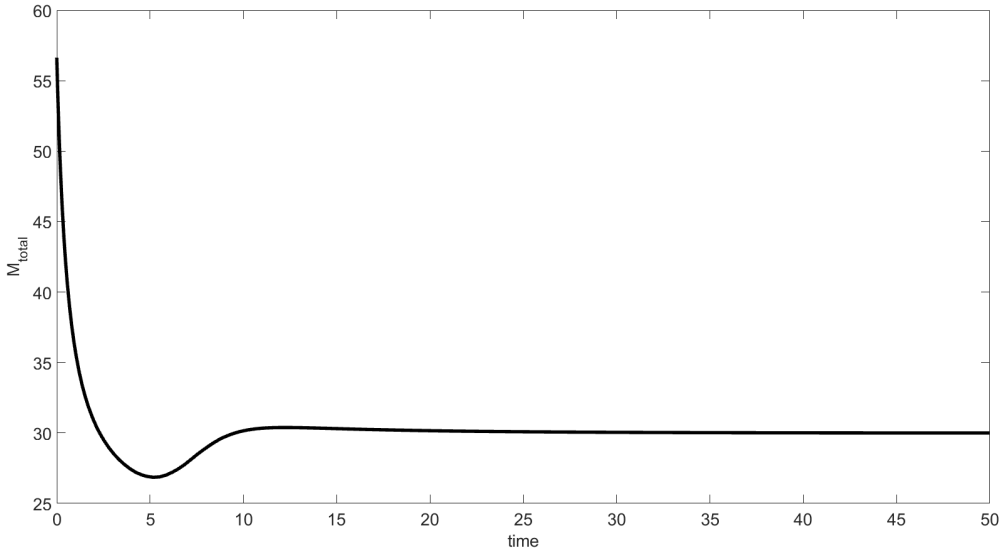
Here we will only consider extreme cases in response to treatment. First resistant parasites are assumed to be fully resistant. Secondly we will examine both cases where the heterozygous parasites act as either resistant or susceptible in response to treatment. However to maintain comparability we set the heterozygous parasites competitive ability at an arbitrary value between the two for both scenarios The parameters that are used in these control examples are given in table 6.2.

As for the Haploid models we first choose to look at the model response to a sustained treatment. Figures 6-4 and 6-5 show this for $\gamma_2 = 1$ (heterozygous parasites being susceptible) and

$\gamma_2 = 0$ (heterozygous are resistant) respectively. In these figures we see that the model has very similar properties to the diploid mean-field model when the heterozygous parasites are susceptible to treatment, namely the initial drop before a slight increase. If the heterozygous parasites act as resistant parasites then this increase does not occur, however the treatment is still less effective than when there is a single susceptible genotype.

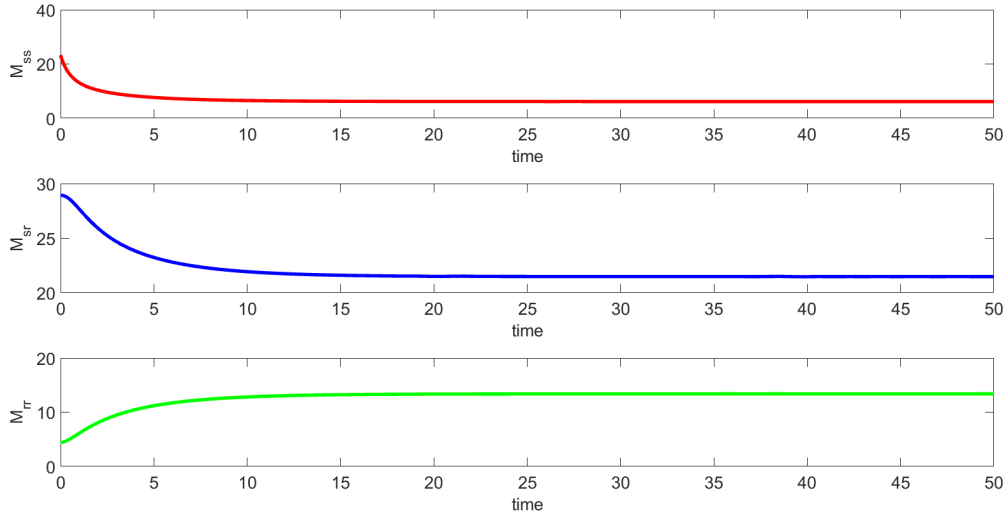


(a)

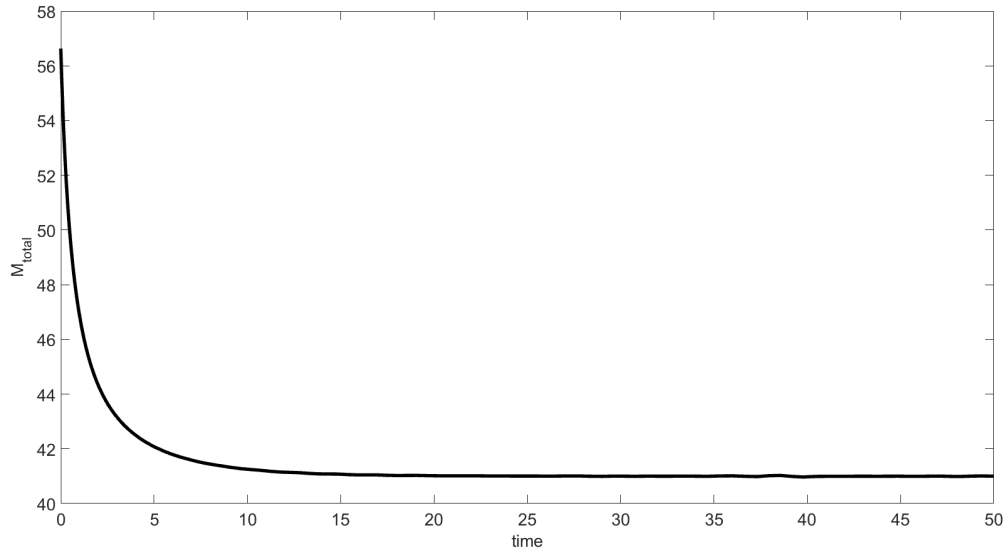


(b)

Figure 6-4: Effects of full intensity sustained treatment on the Diploid mean-field model for $\gamma_2 = 1$, with parameters given in table 6.2.



(a)



(b)

Figure 6-5: Effects of full intensity sustained treatment on the Diploid mean-field model for $\gamma_2 = 0$, with parameters given in table 6.2.

6.2 Continuous Optimal Control

The continuous controls for these models are determined in a very similar way to for the single genotype model shown in Chapter 4. Having defined the dynamics under treatment for each model above we now set up cost functions. While these cost functions take a similar form to the cost functions (4.5) and (4.10), we base cost of the system state on the the total mean and sum of the variances, rather than on the individual genotype means and variances. Additionally we consider actively penalising a high proportion of resistance in the population. We begin by formulating the optimal control problem for the Haploid models, and then we consider the Diploid model.

6.2.1 Haploid Models

For the Haploid models the total mean for both models, mean-field and variance inclusive, is given by:

$$\frac{1}{N}\sum_{f=1}^{N}[M_{i,S}+M_{i,R}] = M_S + M_R.$$

In the variance inclusive Haploid model the total variance is given by:

$$\begin{aligned} V_{total} &= \frac{1}{N}\sum_{i=1}^{N}(M_{i,S}+M_{i,R}-M_S-M_R)^2 \\ &= \frac{1}{N}\sum_{i=1}^{N}\left((M_{i,S}-M_S)^2+(M_{i,R}-M_R)^2+M_{i,S}M_{i,R}-M_SM_{i,R}-M_RM_{i,S}+M_SM_R\right) \\ &= V_S + V_R + C_{S,R} \\ &= V_S + V_R, \end{aligned}$$

where we set $C_{S,R}=0$ as a result of the assumed independence of the two genotypes, discussed in chapter 5. Using these expressions we define the cost functions:

$$J(u) = \int_0^{T_f} u^2 + \alpha_1(M_S+M_R-M_T)^2 dt \tag{6.4}$$

for the mean-field model, and

$$J(u) = \int_0^{T_f} u^2 + \alpha_1(M_S+M_R-M_T)^2 + \alpha_2(V_S+V_R-V_T)^2 dt \tag{6.5}$$

for the variance inclusive model, where α_1, α_2 are positive constants. With these two cost functions the only way that high resistance is penalised is through the potential increase in the mean during sustained treatment and the reduced effectiveness of the treatment itself. As mentioned previously we may wish to more directly penalise high resistance, which we do by considering the proportion of the total population which are resistant to treatment, although this simplifies to become the proportion of the mean population resistant to treatment. An example of this is given in the cost

function:

$$J(u, M_S, M_R) = \int_0^{T_f} u^2 + \alpha_1(M_S + M_R - M_T)^2 + \alpha_r \left(\frac{M_R}{M_R + M_S} - p_T \right)^2 dt, \tag{6.6}$$

for the haploid mean-field model, and

$$J(u, M_S, M_R, V_S, V_R) = \int_0^{T_f} u^2 + \alpha_1(M_S + M_R - M_T)^2 + \alpha_2(V_S + V_R - V_T)^2 + \alpha_r \left(\frac{M_R}{M_R + M_S} - p_T \right)^2 dt, \tag{6.7}$$

for the haploid variance inclusive model.

Mean-Field Model

Combined with the dynamics given in equations (6.1) we derive the Hamiltonian of the haploid mean-field system as:

$$\begin{aligned}
\mathcal{H}(u, M_S, M_R, \lambda_1, \lambda_2) = u^2 + +\alpha_1(M_S + M_R - M_T)^2 + \alpha_r \left(\frac{M_R}{M_R + M_S} - p_T \right)^2 \\
+ \lambda_1 \left(\beta_2 M_S - (\mu_M + u) M_S - D_M(M_S^2 + (1 - \psi) M_S M_R) \right) \\
+ \lambda_2 \left(\beta_2 M_R - \mu_M M_R - D_M(M_R^2 + \psi M_S M_R) \right),
\end{aligned}$$

using the cost function from equation (6.6). The auxiliary variables, λ_1 and λ_2, satisfy

$$\begin{aligned}
\frac{d\lambda_1}{dt} = & -2\alpha_1(M_S + M_R - M_T) + 2\alpha_r \left(\frac{M_R}{(M_R + M_S)^2} \right) \left(\frac{M_R}{M_R + M_S} - p_T \right) \\
& - \lambda_1 \left(\beta_2 - (\mu_M + u) - D_M(2M_S + (1 - \psi) M_R) \right) \\
& - \lambda_2 \left(-D_M \psi M_R \right)
\end{aligned}$$

$$\begin{aligned}
\frac{d\lambda_2}{dt} = & -2\alpha_1(M_S + M_R - M_T) - 2\alpha_r \left(\frac{M_S}{(M_R + M_S)^2} \right) \left(\frac{M_R}{M_R + M_S} - p_T \right) \\
& - \lambda_1 \left(-D_M(1 - \psi) M_S \right) \\
& - \lambda_2 \left(\beta_2 - \mu_M - D_M(2M_R) + \psi M_S \right),
\end{aligned}$$

which have the conditions that $\lambda_1(T_f) = \lambda_2(T_f) = 0$ as the terminal state is unfixed. Following the same methodology as section 4.1 the condition of optimality is given by

$$\frac{\partial \mathcal{H}(u^*, M_S^*, M_R^*, \lambda_1^*, \lambda_2^*)}{\partial u} = 2u^* - \lambda_1 M_S = 0,$$

Parameter set	Parameter	Value
Set 1	M_T	20.64
	p_T	0.10
	α_1	$\frac{1}{6.23^2}$
	α_r	0
Set 2	M_T	20.64
	p_T	0.10
	α_1	$\frac{1}{2(6.23^2)}$
	α_r	$\frac{1}{2(0.32^2)}$

Table 6.3: Table showing two parameter sets used for the cost function (6.6) where set 1 only aims to minimise the total mean, set 2 aims to minimise a combination of the total mean and the proportion of resistant parasites in population

when $0 \leq u^* \leq 1$. This gives the value of u^* as

$$u^* = \begin{cases} 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0 \\ \frac{\lambda_1 M_S}{2} & \text{if } \frac{\partial \mathcal{H}}{\partial u} = 0. \end{cases}$$

The target values, M_T and p_T, are set to the minimum values of M_{total} and $\frac{M_R}{M_R+M_S}$ that occur when a treatment is applied to a previously untreated population. These target values correspond to the minimum values shown, for example in Figure 6-1. The values of α_1 and α_r are set as in Chapter 4 (see section 4.1.3) and are given as

$$\alpha_1 = \frac{q}{(M_x - M_T)^2} \quad \alpha_r = \frac{1-q}{(p_x - p_T)^2}.$$

If we do not wish to directly penalise the proportion of resistance in the population we set $q = 1$, and $\alpha_r = 0$, otherwise we may set $0 < q < 1$. For the examples we present in this section the parameters that we use are shown in table 6.3.

Figure 6-6 shows the optimal continuous controls as determined using the BOCOP software for the two parameter sets for the cost functions in this table, and the model parameters shown in table 6.3. We see very different properties to the optimal controls for the basic mean-field model. The control here starts out low, and remains at a a relatively low level, compared with the single genotype models, until the end when it suddenly increases. This happens because treating the hosts early leads to a reduction in the susceptible population which reduces the density dependent mortality of resistant parasites which allows that population to increase. Control may be applied to the end of the period of interest as the period ends before the effects of the increased cost of resistance are felt.

When the proportion of resistance is not directly penalised and $\alpha_r = 0$, as in figure 6-6 sub-figures (a)/(b), we can see that the control favours a swifter increase than when the proportion of

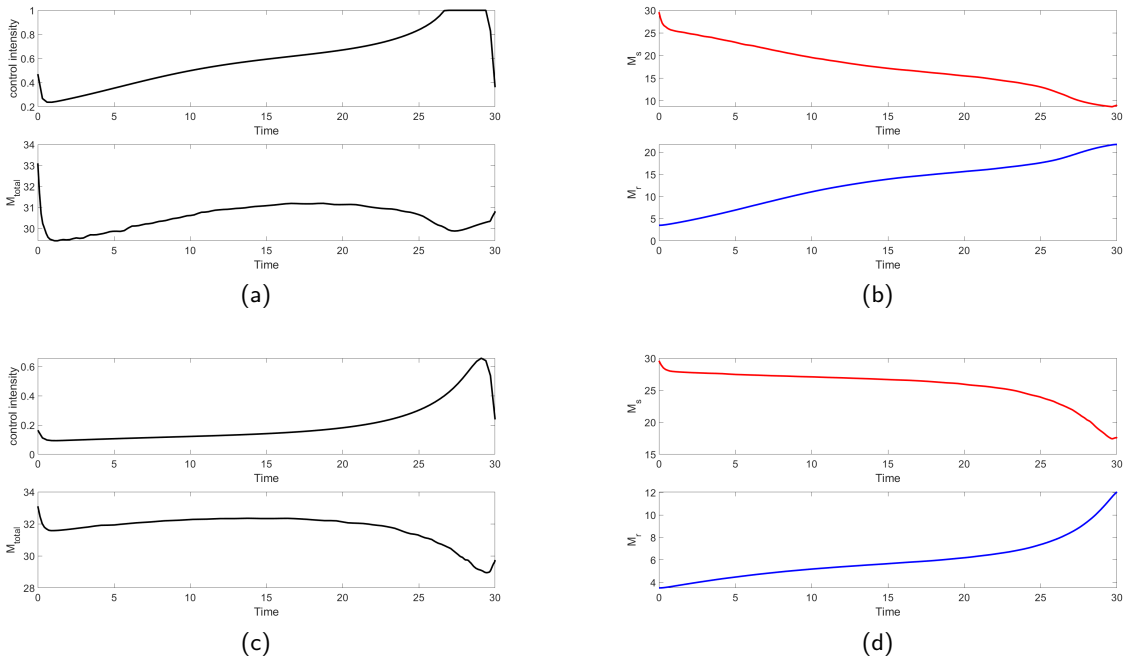


Figure 6-6: Continuous optimal control on haploid mean-field model (6.1) with parameters for dynamics in table 6.1 and cost function parameters given by (a)/(b) Set 1 (c)/(d) Set 2 in table 6.3

resistance is directly accounted for, $\alpha_r > 0$ (sub-figures (c)/(d)) as is expected. It is interesting that the optimal control does not seem to take much advantage of the reactivity of the total mean and apply greater control for a short initial period to maximise the reduction before resistance builds up. This suggests that the benefit of the initial reduction is outweighed by the increased cost that follows. How this will affect controls which lack the ability to look ahead, such as the sequential integer controls will be discussed in the next chapter.

Variance Inclusive Model

For the variance inclusive haploid model we use the cost function (6.7). Formulating the Hamiltonian from this and the dynamics given in (6.2) gives:

$$\begin{aligned}\mathcal{H}(u, M_S, M_R, \lambda_1, \lambda_2) = u^2 + &+\alpha_1(M_S + M_R - M_T)^2 + \alpha_r \left(\frac{M_R}{M_R + M_S} - p_T\right)^2 \\ &+ \lambda_1 \left(\beta_2 M_S - (\mu_M + u)M_S - D_M(M_S^2 + V_S + (1-\psi)M_S M_R)\right) \\ &+ \lambda_2 \left(\beta_2 M_R - \mu_M M_R - D_M(M_R^2 + V_R + \psi M_S M_R)\right). \\ &- \lambda_3 \left(\beta_2 M_S - (\mu_M + u)(2V_S - M_S)\right. \\ &\left. -D_M \left(\frac{4V_S^2}{M_S} + 4M_S V_S + M_S^2 - 3V_S + (1-\psi)(2M_R V_S - M_R M_S)\right)\right) \\ &+ \lambda_4 \left(\beta_2 M_R - (\mu_M + \gamma u)(2V_R - M_R)\right. \\ &\left. -D_M \left(\frac{4V_R^2}{M_R} + 4M_R V_R - M_R^2 - 3V_R + \psi(2M_S V_R - M_S M_R)\right)\right).\end{aligned}$$

Taking the partial derivatives of this with respect to M_S, M_R, V_S and V_R respectively give that the auxiliary variables will satisfy the following ODES

$$\begin{aligned}\frac{d\lambda_1}{dt} &= -2\alpha_1(M_S + M_R - M_T) + 2\alpha_r \left(\frac{M_R}{(M_R + M_S)^2}\right)\left(\frac{M_R}{M_R + M_S} - p_T\right) \\ &\quad - \lambda_1 \left(\beta_2 - (\mu_M + u) - D_M(2M_S + (1-\psi)M_R)\right) \\ &\quad - \lambda_2 \left(-D_M \psi M_R\right) \\ &\quad - \lambda_3 \left(\beta_2 + \mu_M + u - D_M \left(\frac{-4V_S^2}{M_S^2} + 4V_S - 2M_S - (1-\psi)M_R\right)\right) \\ &\quad - \lambda_4 \left(-D_M \psi (V_R + M_R)\right) \\ \frac{d\lambda_2}{dt} &= -2\alpha_1(M_S + M_R - M_T) - 2\alpha_r \left(\frac{M_S}{(M_R + M_S)^2}\right)\left(\frac{M_R}{M_R + M_S} - p_T\right) \\ &\quad - \lambda_1 \left(-D_M(1-\psi)M_S\right) \\ &\quad - \lambda_2 \left(\beta_2 - \mu_M - D_M(2M_R) + \psi M_S\right) \\ &\quad - \lambda_3 \left(-D_M(1-\psi)(V_S + M_S)\right) \\ &\quad - \lambda_4 \left(\beta_2 + \mu_M - D_M \left(\frac{-4V_R^2}{M_R^2} + 4V_R - 2M_R - \psi M_S\right)\right) \\ \frac{d\lambda_3}{dt} &= -2\alpha_2(V_S + V_R - V_T) + \lambda_1 D_M \\ &\quad - \lambda_3 \left(-2(\mu_M + u) - D_M \left(\frac{8V_S}{M_S} + 4M_S - 3 + 2(1-\psi)M_R\right)\right) \\ \frac{d\lambda_4}{dt} &= -2\alpha_2(V_S + V_R - V_T) + \lambda_2 D_M \\ &\quad - \lambda_4 \left(-2\mu_M - D_M \left(\frac{8V_R}{M_R} + 4M_R - 3 + 2\psi M_S\right)\right),\end{aligned}$$

with terminal conditions $\lambda_1(T_f) = \lambda_2(T_f) = \lambda_3(T_f) = \lambda_4(T_f) = 0$. The condition of optimality is given by

$$\frac{\partial \mathcal{H}}{\partial u} = 2u - \lambda_1 M_S - \lambda_3(2V_S - M_S) = 0.$$

Combined with the conditions we get due to the boundedness of u the optimal control is given by

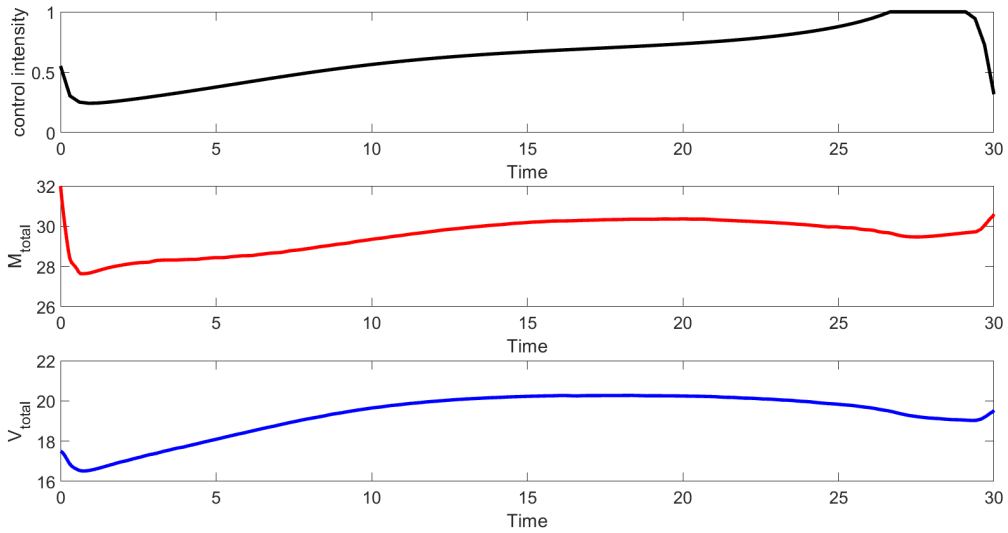
$$u^* = \begin{cases} 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0 \\ \frac{\lambda_1 M_S + \lambda_2(2V_S - M_S)}{2} & \text{if } \frac{\mathcal{H}}{\partial u} = 0. \end{cases}$$

Using the same method of determining the cost function parameters as for the basic model and the mean-field model shown above, we set the parameters $M_T,\ V_T,\ p_T,\ \alpha_1,\ \alpha_2$ and α_r for our examples as in table 6.4. These parameter sets aim to minimise either the mean, the variance, or a combination of both with the maximal acceptable value of the total mean and variance (M_x and V_x) used to set up the parameters set halfway between the untreated equilibrium value and the target values. This gives $M_x = \frac{M_T + M^*_{total}}{2}$ and $V_x = \frac{V_T + V^*_{total}}{2}$, where M^*_{total} and V^*_{total} are the untreated equilibrium values. We also include parameter sets which directly try to minimise the proportion of resistant parasites in the host, just as we did for the mean-field model.

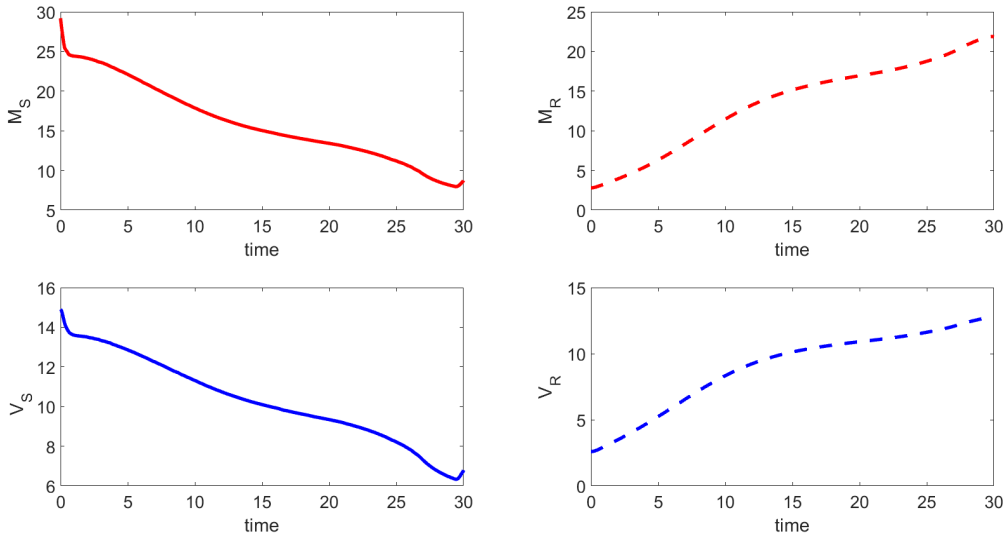
Figures 6-7 and 6-8 show examples of continuous optimal control for parameter sets 1 and 3 given in table 6.4. Studying these figures we see that when the control aims to minimise the mean or the mean and the proportion of resistance as it does in figure 6-7 the controls are very similar to those given by the haploid mean-field model, shown previously, and start far lower and gradually increase over the interval before a short period of full intensity right before the end. Figures 6-8 show an example of the optimal control when the aim is to minimise the variance towards it target value. The controls we see here are interesting in how much they differ from those determined by minimising the mean. In these examples an initial high intensity control is applied which drops both the mean and the variance swiftly. After this initial burst of control the controls are kept at a very similar level, which is high when compared with the control on the mean, and then a final high intensity application is done right at the end. What is interesting about these controls is that for a large amount of time over the interval the variance is actually at a higher level than if treatment had not been applied at all. This suggests that the initial drop in the variance was great enough to outweigh the cost of the control and the increased variance that results from this treatment later in the interval. Unlike what we observed for the controls based on the mean this control takes far more advantage of the initial reactivity and the reduction that this brings in the total variance. In general we found that controls which aimed to reduce the mean behaved as would be expected, with the majority of control applied towards the end of the control period where the subsequent increased cost would not come into effect during the time period being examined. Controls based on the variance acted in a more unexpected manner, with an abundance of control applied initially and the result being the variance increasing after the initial drop. This was followed by a further burst of high intensity treatment at the end of the time period, similar to controls based on the

Parameter set	Parameter	Value	Parameter set	Parameter	Value
Set 1	M_T	19.0733	Set 2	M_T	19.0733
	V_T	14.0783		V_T	14.0783
	p_T	0.0873		p_T	0.0873
	α_1	$\frac{1}{6.4592^2}$		α_1	$\frac{1}{2(6.4592^2)}$
	α_2	0		α_2	0
	α_r	0		α_r	$\frac{1}{2(0.3427^2)}$
Set 3	M_T	19.0733	Set 4	M_T	19.0733
	V_T	14.0783		V_T	14.0783
	p_T	0.0873		p_T	0.0873
	α_1	0		α_1	0
	α_2	$\frac{1}{1.6954^2}$		α_2	$\frac{1}{2(1.6954^2)}$
	α_r	0		α_r	$\frac{1}{2(0.3427^2)}$
Set 5	M_T	19.0733	Set 6	M_T	19.0733
	V_T	14.0783		V_T	14.0783
	p_T	0.0873		p_T	0.0873
	α_1	$\frac{1}{2(6.4592^2)}$		α_1	$\frac{1}{3(6.4592^2)}$
	α_2	$\frac{1}{2(1.6954^2)}$		α_2	$\frac{1}{3(1.6954^2)}$
	α_r	0		α_r	$\frac{1}{3(0.3427^2)}$

Table 6.4: Parameter sets for the examples of control on the variance inclusive haploid model. Sets 1 and 2 aim to minimise the mean, sets 3 and 4 the variance and sets 5 and 6 a combination of both. The even numbered parameter sets include a non-zero α_r parameter to directly include the minimisation of the proportion of resistance.

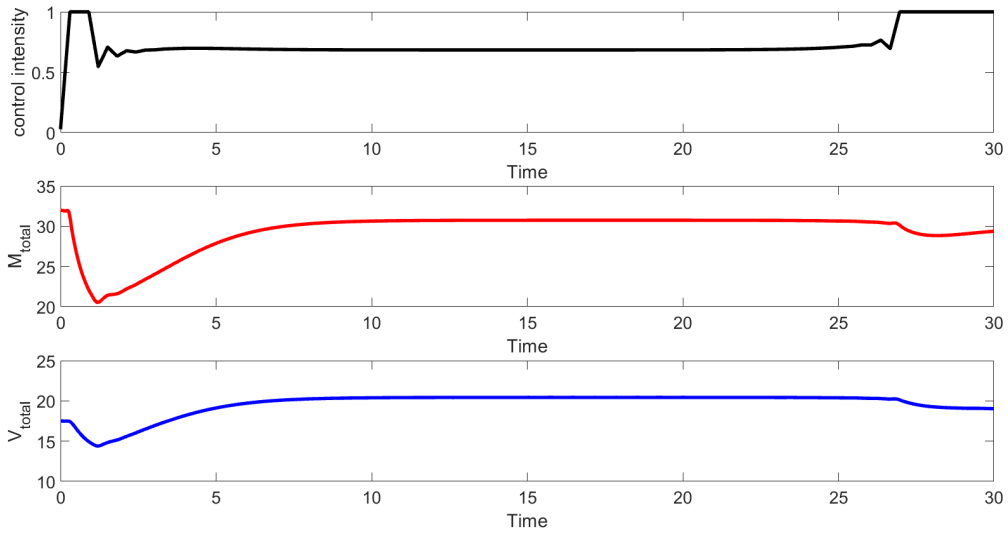


(a)

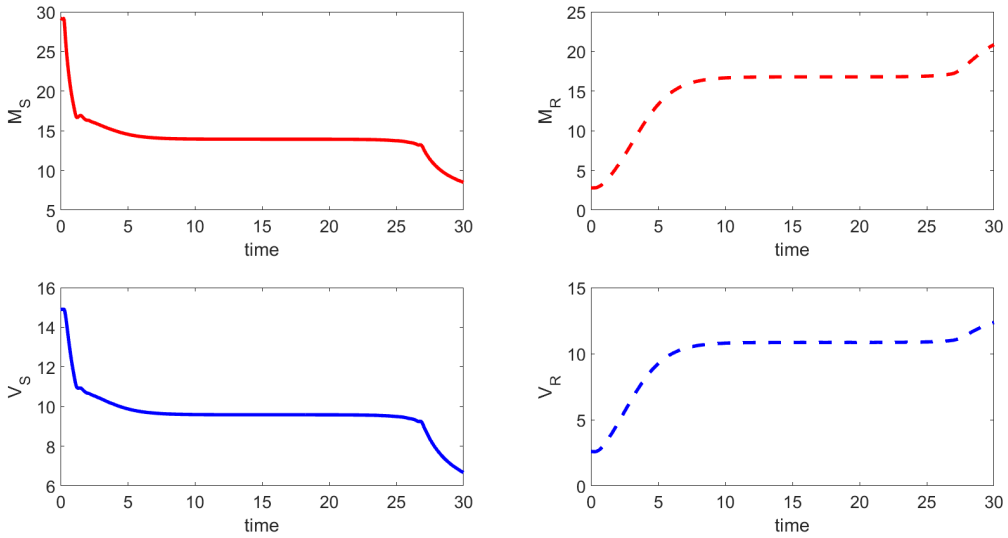


(b)

Figure 6-7: Continuous optimal control on haploid variance inclusive model (6.2) with parameters for dynamics in table 6.1 and cost function parameters given by Set 1 in table 6.4.



(a)



(b)

Figure 6-8: Continuous optimal control on haploid variance inclusive model (6.2) with parameters for dynamics in table 6.1 and cost function parameters given by Set 3 in table 6.4.

mean. These results suggest that controls based on the variance try to take advantage of the initial reduction that may be achieved, in spite of the dynamics leading to a less favourable state later on in the time period, while controls based on the mean do not.

6.2.2 Diploid Model

Given the dynamics under treatment defined in equation (6.3) then we simply need a cost function in order to set up the optimal control problem for the diploid mean-field model. For this model we may set up an equation similar to those in equations (6.6) and (6.7), although further thought is required to include a term to minimise the proportion of resistance. This is once again due to the presence of the heterozygous parasites. The most obvious option may be to set the cost function to

$$J(u) = \int_0^{T_F} u^2 + \alpha_1 (M_{SS} + M_{SR} + M_{RR} - M_T)^2 + \alpha_r \left(\frac{2M_{RR} + M_{SR}}{2(M_{SS} + M_{SR} + M_{RR})} - p_T \right)^2 dt. \quad (6.8)$$

Here the proportion of resistance is measuring the proportion of resistant alleles within the population rather than how many parasites are truly resistant. Another option is, if the heterozygous parasites act as either fully resistant or fully susceptible phenotypes, to measure the proportion of parasites which act as fully resistant. This would be given by either

$$\frac{M_{RR} + M_{SR}}{(M_{SS} + M_{SR} + M_{RR})} \qquad \text{or} \qquad \frac{M_{RR}}{(M_{SS} + M_{SR} + M_{RR})},$$

if the heterozygous parasites act as resistant or susceptible, respectively. For the purposes of control here we will use (6.8) and the proportion of resistant alleles where applicable.

Setting up the Hamiltonian system for this model using (6.3) and (6.8) gives:

$$\begin{aligned}
&\mathcal{H}(u, M_{SS}, M_{SR}, M_{RR}, \lambda_1, \lambda_2, \lambda_3) \\
&= u^2 + \alpha_1 (M_{SS} + M_{SR} + M_{RR} - M_T)^2 + \alpha_r \left(\frac{2M_{RR} + M_{SR}}{2(M_{SS} + M_{SR} + M_{RR})} - p_T \right)^2 \\
&+ \lambda_1 \left(\beta_2 \frac{M_{SS}^2 + M_{SS} M_{SR} + 0.25 M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - (\mu_M + u) M_{SS} \right. \\
&\left. - D_M M_{SS} (M_{SS} + (1 - \psi_1) M_{SR} + (1 - \psi_2) M_{RR}) \right) \\
&+ \lambda_2 \left(\beta_2 \frac{2M_{SS} M_{RR} + M_{SS} M_{SR} + 0.5 M_{SR}^2 + M_{RR} M_{SR}}{M_{SS} + M_{SR} + M_{RR}} - (\mu_M + \gamma_2 u) M_{SR} \right. \\
&\left. - D_M M_{SR} (M_{SR} + (\psi_1) M_{SR} + (1 - \psi_3) M_{RR}) \right) \\
&+ \lambda_3 \left(\beta_2 \frac{M_{RR}^2 + M_{RR} M_{SR} + 0.25 M_{SR}^2}{M_{SS} + M_{SR} + M_{RR}} - (\mu_M + \gamma_1 u) M_{RR} \right. \\
&\left. - D_M M_{RR} (M_{RR} + (\psi_2) M_{SS} + (\psi_3) M_{SR}) \right).
\end{aligned} \quad (6.9)$$

Our auxiliary functions then satisfy the following ODES:

$$\begin{aligned}
\frac{d\lambda_1}{dt} &= -2\alpha_2(M_{SS}+M_{SR}+M_{RR}-M_T) \\
&- 2\alpha_r\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})^2}\right)\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})}-p_T\right) \\
&- \lambda_1\left(\beta_2\left(\frac{2M_{SS}+M_{SR}}{M_{SS}+M_{SR}+M_{RR}}-\frac{M_{SS}^2+M_{SS}M_{SR}+0.25M_{SR}^2}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-(\mu_M+u\right. \\
&\left.-D_M(2M_{SS}+(1-\psi_1)M_{SR}+(1-\psi_2)M_{RR})\right) \\
&- \lambda_2\left(\beta_2\left(\frac{2M_{RR}+M_{SR}}{M_{SS}+M_{SR}+M_{RR}}-\frac{2M_{SS}M_{RR}+M_{SS}M_{SR}+0.5M_{SR}^2+M_{RR}M_{SR}}{(M_{SS}+M_{SR}+M_{RR})^2}\right)\right. \\
&\left.-D_M(\psi_1 M_{SR})\right) \\
&- \lambda_3\left(-\beta_2\left(\frac{M_{RR}^2+M_{RR}M_{SR}+0.25M_{SR}^2}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-D_M(\psi_2 M_{RR})\right)
\end{aligned}$$

$$\begin{aligned}
\frac{d\lambda_2}{dt} &= -2\alpha_2(M_{SS}+M_{SR}+M_{RR}-M_T) \\
&- \alpha_r\left(2\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})^2}\right)-\frac{1}{2(M_{SS}+M_{SR}+M_{RR})}\right)\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})}-p_T\right) \\
&- \lambda_1\left(\beta_2\left(\frac{M_{SS}+0.5M_{SR}}{M_{SS}+M_{SR}+M_{RR}}-\frac{M_{SS}^2+M_{SS}M_{SR}+0.25M_{SR}^2}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-D_M\psi_1 M_{SS}\right) \\
&- \lambda_2\left(\beta_2\left(1-\frac{2M_{SS}M_{RR}+M_{SS}M_{SR}+0.5M_{SR}^2+M_{RR}M_{SR}}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-(\mu_M+\gamma_2 u\right. \\
&\left.-D_M(2M_{SR}+\psi_1 M_{SS}+(1-\psi_3)M_{RR})\right)
\end{aligned}$$

$$\begin{aligned}
\frac{d\lambda_3}{dt} &= -2\alpha_2(M_{SS}+M_{SR}+M_{RR}-M_T) \\
&- \alpha_r\left(2\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})^2}\right)-\frac{1}{M_{SS}+M_{SR}+M_{RR}}\right)\left(\frac{2M_{RR}+M_{SR}}{2(M_{SS}+M_{SR}+M_{RR})}-p_T\right) \\
&- \lambda_1\left(-\beta_2\left(\frac{M_{SS}^2+M_{SS}M_{SR}+0.25M_{SR}^2}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-D_M((1-\psi_2)M_{SS})\right) \\
&- \lambda_2\left(\beta_2\left(\frac{2M_{SS}+M_{SR}}{M_{SS}+M_{SR}+M_{RR}}-\frac{2M_{SS}M_{RR}+M_{SS}M_{SR}+0.5M_{SR}^2+M_{RR}M_{SR}}{(M_{SS}+M_{SR}+M_{RR})^2}\right)\right. \\
&\left.-D_M(\psi_3 M_{SR})\right) \\
&- \lambda_3\left(\beta_2\left(\frac{2M_{RR}+M_{SR}}{M_{SS}+M_{SR}+M_{RR}}-\frac{M_{RR}^2+M_{RR}M_{SR}+0.25M_{SR}^2}{(M_{SS}+M_{SR}+M_{RR})^2}\right)-(\mu_M+u\right. \\
&\left.-D_M(2M_{RR}+\psi_3 M_{SR}+\psi_2 M_{SS})\right),
\end{aligned}$$

with the terminal conditions once again given by $\lambda_1(T_f) = \lambda_2(T_f) = \lambda_3(T_f) = 0$. Partially differentiating the Hamiltonian with respect to u gives the optimality condition :

$$\frac{\partial \mathcal{H}}{\partial u} = 2u - \lambda_1 M_{SS} - \lambda_2\gamma_2 M_{SR} - \lambda\gamma_1 M_{RR} = 0.$$

Parameter set	Parameter	Value
Set 1	M_T	26.8537
	p_T	0.3334
	α_1	$\frac{1}{14.8894^2}$
	α_r	0
Set 2	M_T	26.8537
	p_T	0.1162
	α_1	$\frac{1}{2(14.8894^2)}$
	α_r	$\frac{1}{2(0.3331^2)}$

Table 6.5: Cost function parameters for the diploid cost function (6.8) when the heterozygous parasites are susceptible and $\gamma_2 = 1$, for use with dynamic parameters given in table 6.2

.

When used in conjunction with the conditions that result from the boundedness of u we get that

$$u^* = \begin{cases} 0 & \text{if } \frac{\partial \mathcal{H}}{\partial u} > 0 \\ 1 & \text{if } \frac{\partial \mathcal{H}}{\partial u} < 0 \\ \frac{\lambda_1 M_{SS} + \lambda_2 \gamma_2 M_{SR} + \lambda_3 \gamma_1 M_{RR}}{2} & \text{if } \frac{\partial \mathcal{H}}{\partial u} = 0. \end{cases}$$

As we have done for the haploid models we set parameters for the cost functions by basing them on the minimal achievable total mean, and proportion of resistance, assuming a full intensity treatment is applied to equation (6.3) with initial conditions set at the untreated equilibrium values. The cost function parameters α_1 and α_r are set using the same method as before, depending on whether the proportion of resistance is being directly penalised or not. As we wish to study cases where the heterozygous parasites act as either susceptible or resistant parasites we give different cost function parameter sets for each of these cases, as the minimum achievable mean will be different in each case. These parameter values are given in tables 6.5, when the heterozygous parasites act as susceptible parasites and $\gamma_2 = 1$, and table 6.6, for examples in which the SR genotype parasites act as resistant parasites and $\gamma_2 = 0$.

Figure 6-9 shows the optimal controls determined on the diploid model given in (6.3) using the cost function parameters given in table 6.5 and the corresponding model parameters from table 6.2 with $\gamma_2 = 1$. The control in these figures bears similarity to the controls determined by the basic model. The control starts high before reducing to a more moderate level for the majority of the interval. The initial control leads to a reduction in the fully susceptible and heterozygous parasites and a minor increase in the resistant parasites. This increase in the resistant parasites that results from the lack of competition has limited effect as the reduction in the heterozygous parasites also reduces the birth rate of the resistant parasites. This reduced birth rate results in the resistant parasite population reducing again, despite not being affected by the treatment directly. When the proportion of resistance is minimised as well as the mean, as in sub-figures (b) and (d), the intensity

Parameter set	Parameter	Value
Set 3	M_T	40.9683
	p_T	0.3334
	α_1	$\frac{1}{7.8321^2}$
	α_r	0
Set 4	M_T	40.9683
	p_T	0.3334
	α_1	$\frac{1}{2(7.8321^2)}$
	α_r	$\frac{1}{2(0.1275^2)}$

Table 6.6: Cost function parameters for the diploid cost function (6.8) when the heterozygous parasites are resistant and $\gamma_2 = 0$, for use with dynamic parameters given in table 6.2

.

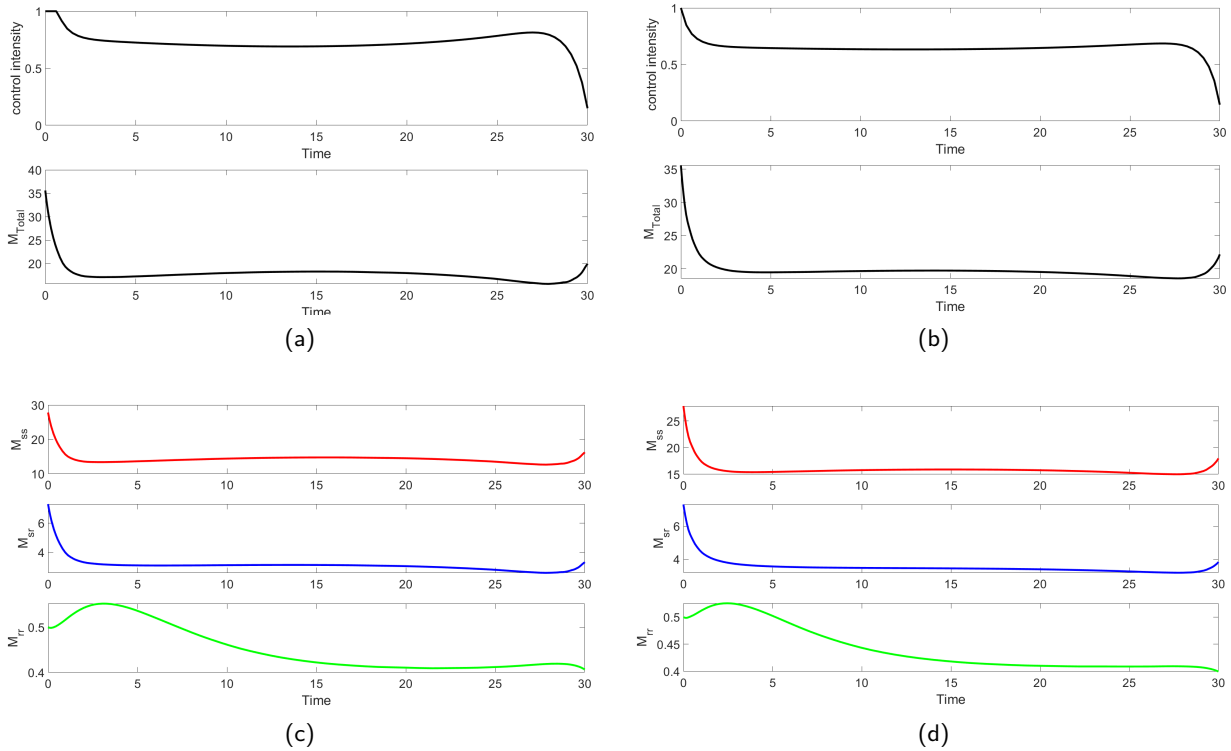


Figure 6-9: Continuous optimal control on diploid mean-field model (6.3) with parameters for dynamics in table 6.2, with $\gamma_2 = 1$, and cost function parameters given by (a) and (c) Set 1 in table 6.5 (b) and (d) Set 2 in table 6.5

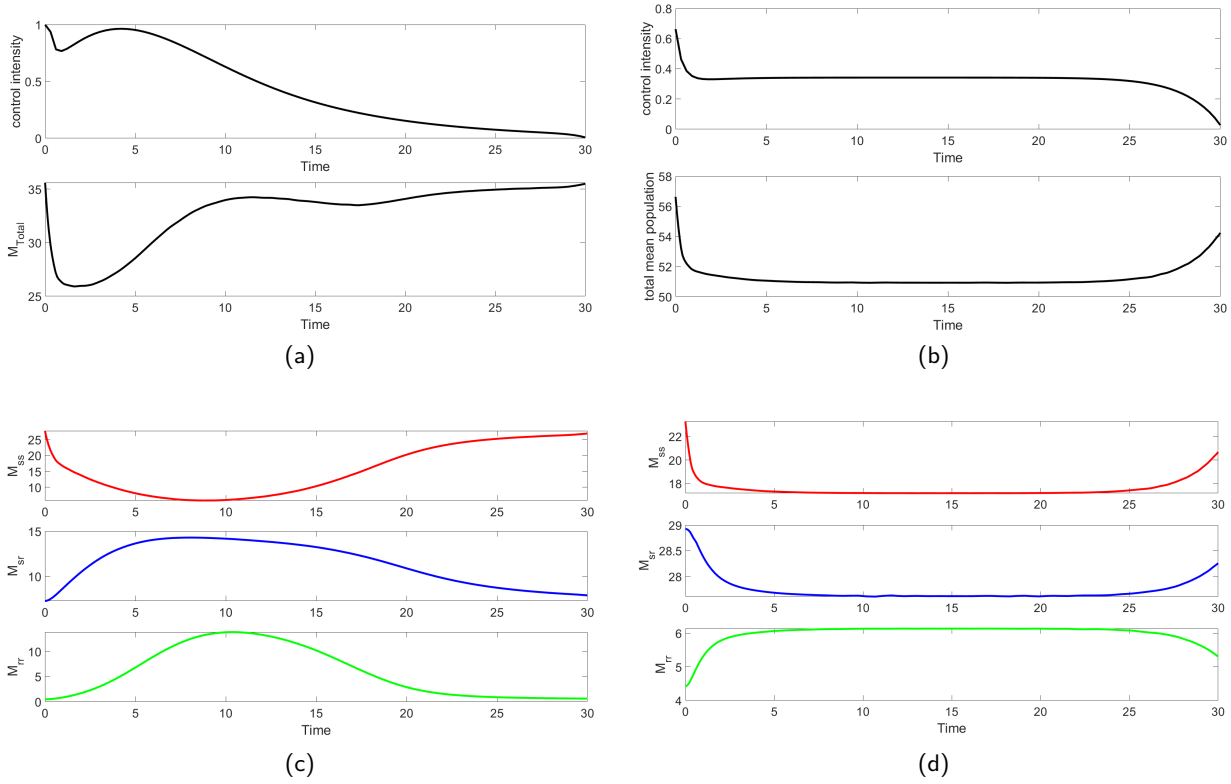


Figure 6-10: Continuous optimal control on diploid mean-field model (6.3) with parameters for dynamics in table 6.2, with $\gamma_2 = 0$, and cost function parameters given by (a) and (c) Set 3 in table 6.5 (b) and (d) Set 4 in table 6.5

of the treatment is reduced far sooner than when it is not directly accounted for, but the treatment following this is qualitatively very similar. This shorter period of maximal intensity has the result of reducing the susceptible and heterozygous parasites to a lesser extent than the treatment shown in sub-figures (a) and (c). This prevents the population of resistant parasites increasing as much, and keeps the proportion of resistance lower. The sustained treatment at the lower level then maintains the reduction of all genotypes that followed the initial peak in the resistant parasites. While the control appears similar to that on the single genotype mean-field model, the effects that these controls have on the populations are notable different. In this case, we see that the initial treatment reduces the mean quite significantly in both cases but, after the initial drop, the total mean undergoes a small increase again.

As a comparison, figure 6-10 shows the controls determined for a model with resistant heterozygous parasites, that is $\gamma_2 = 0$, with the cost function parameters given in sets 3 and 4 in table 6.6. The parameters for the underlying dynamics are kept the same. When we set our cost function parameters to reduce the mean (parameter set 3), as in figure 6-10 sub-figures (a) and (c), the control starts high but is rapidly reduced. The total mean behaves as would be expected in this situation with an initial drop corresponding to the application of treatment before increasing to the untreated equilibrium once treatment is stopped. When the proportion of resistance is also minimised (parameter set 4, figure 6-10 sub-figures (b) and (d)) the control is actually applied for far longer. Common to both of these simulations is that, despite being resistant to control, the population of heterozygous parasites actually decreases when treatment is applied. This is due to the decrease in the birth rate of heterozygous parasites caused by a lack of susceptible parasites. Comparing these it is also notable that a control that aims to reduce the proportion leads to more control, which is the opposite of the results seen in the diploid model. This suggests two things; firstly, the reduction in heterozygous parasites actually lowers the proportion of resistant alleles in the population more than the increase in the resistant parasite increases it. Secondly, the benefit of reducing the proportion is greater than the benefit of reducing the total mean, at least for this particular choice of cost function parameters.

6.3 Discussion

What this section has shown is that when resistance is included in the model as a genetic trait, which is tracked, the response to treatment is quite different. We saw in figures 6-1, 6-3, and 6-4 that under the application of treatment the state variables tend to their equilibrium values with little reactivity. However, due to the different rates at which they do this when we consider what happens to the summations of the state variables, M_{total} and V_{total}, we see that the total mean and variance can actually have an initial drop before they increase to the value they take at the treated equilibria depending on the parameters. This bears similarity to the transient growth property that may be seen when the model is reactive except that it is not the individual variables showing transient growth

but the summations of the variables.

The other crucial difference that we found when investigating this is that when we compared the optimal controls on the two and three genotype models we found that the two genotype model favoured control towards the end of the treatment interval, where the increased cost of the treatment in the future will have no effect, while the three genotype model had optimal controls which applied most control at the start of the interval. This links to the previous property, as when going from the treated equilibrium to the untreated equilibrium the same property is seen. The haploid models increase past the equilibrium initially and then decrease again, while this is far less significant in the diploid model. The result of this is that stopping treatment on the haploid model makes the infection significantly worse than if it had been untreated. To avoid this, it is held off until the end when what occurs later is no longer important.

When it comes to the inclusion of the variance the haploid model did not give markedly different results when it was included. This was not unexpected given the results of the basic model and the lack of covariance between the two populations.

The main result of this chapter is that when a population is made up of interacting sub-groups the response of the individual sub-groups may not be a good indicator of how the overall population will respond to control. Having seen this the following chapter will explore integer treatments on the resistant models, much like for the single genotype models in chapter 4.

Chapter 7

Integer Controls on Resistant Models

On the single genotype models, discussed in chapters 3 and 4, after having determined the continuous optimal controls for each of the mean-field and variance inclusive models, integer controls were determined. For each model, an integer control was determined by choosing the optimal strategy on each sub-interval sequentially. On these controls, the length of the sub-intervals was a defining factor in the behaviour and cost of the control, with shorter intervals seeing less benefit from treatment and thus more likely to avoid treatment. When the control was switched between on and off a dynamic equilibrium was often shown to arise between the model state and the control. The second method by which integer controls were determined was to optimise the treatment on the sub-intervals to minimise the cost across the full interval. On the single genotype models, these optimal integer controls showed similarities to the respective continuous optimal controls, with full intensity treatment being applied until near the end of the interval. This chapter explores the integer controls determined by these same methodologies on the resistance inclusive models.

The properties of our resistant models that have been seen previously make the potential controls that may result from these methods less predictable. In particular, it has been shown that sustained treatment can increase the proportion of resistance in the population, rendering the treatment less effective, figure 6-1, 6-3 and 6-4. This change in the efficiency of treatment could lead to more fluctuation in the control. Whether this is the case is discussed in the remainder of this chapter.

7.1 Haploid Mean-Field Model

As we did for the basic models we first determine a control on the model by splitting the overall time interval that we apply control to into equally sized sub-intervals. With these sub-intervals, we first determine a control by choosing to turn the control on or off for each interval based on whether it is advantageous for that singular interval, given the initial conditions of that individual interval. The initial conditions for each interval are dependent on what action was taken in the previous interval. As for the continuous optimal control the dynamics of the model may be given by (6.1), with the

treatment variable restricted to $u \in \{0, 1\}$. We use the same cost function as for the continuous optimal control, (6.6), to determine the lowest cost option on an interval and apply this. If we use the same parameter set for the dynamics, given in table 6.1, and the same two sets of parameters for the cost function, table 6.3, then we find the examples shown in figure 7-1 and 7-2 where the sub-interval length is changed between 0.25, 0.5 and, 1 for each parameter set.

Studying these examples we see that when the proportion of resistance is not minimised explicitly (table 5.1, parameter set 1), the model state tends towards the treated equilibrium. For longer sub-intervals, as seen in figure 7-1, sub-figures (e)/(f), the treatment is sustained across all intervals forcing the state to the treated equilibrium. For shorter intervals system ultimately reaches approximately the treated equilibrium, but allows for the treatment to be switched off for short periods. If we compare the cost of the controls that we see in figure 7-1 we see the cost breakdown in figure 7-3. This shows that decreasing the sub-interval length can first improve the cost estimate but further decreasing it causes the cost to rise. By looking at the dynamics of the model when treatment is repeatedly stopped and started in quick succession we can see one reason why this effect is perhaps more pronounced. When treatment is stopped, the susceptible parasite population builds up. In the long run, this would lead to a drop in the resistant parasites as a result of competition. When the treatment is only stopped for short periods the reduction in resistant parasites is too slow to have a significant effect so the considerably faster increase in susceptible parasites outweighs the small decrease in resistant parasites. The treatment then gets switched on to counteract this increase in susceptible parasites and, as a result, allows the resistant population parasite to increase further. The result of this can be that in the time following treatment the total mean can increase beyond the value it takes when at the untreated equilibrium. The extent to which this may occur depends on how much the resistant parasite populations have built up during treatment. This is a property that was discussed in the previous chapter in relation to its effects on the optimal controls.

When we adapt the cost function parameters to explicitly minimise the proportion of resistant parasites as well as the mean population size the results showed quite significant change. This inclusion meant that strategies in general favoured less treatment. This is because minimising the proportion of resistance is achieved by not treating, which was the result given in figure 7-2 sub-figures (a)/(b). It is noted that due to the the the rounding of the initial conditions the model is not perfectly at the untreated equilibrium initially so does drop despite the lack of treatment. Determining control using the parameters from set 2 with longer time intervals gives the results in sub-figures (c)/(d) and ((e)/(f). These show treatment which fluctuates between on and off can cause a greater total mean than would be reached if treatment were not applied. In the remaining examples in this chapter, all parameter sets for the cost functions are set to explicitly minimise the proportion of resistance in the population.

Next we considered how the integer control may differ when it is optimised across the entire interval. As we saw in the basic models this puts limitations on the overall length of the interval as it is computationally costly with each sub-interval doubling the number of potential controls. Once

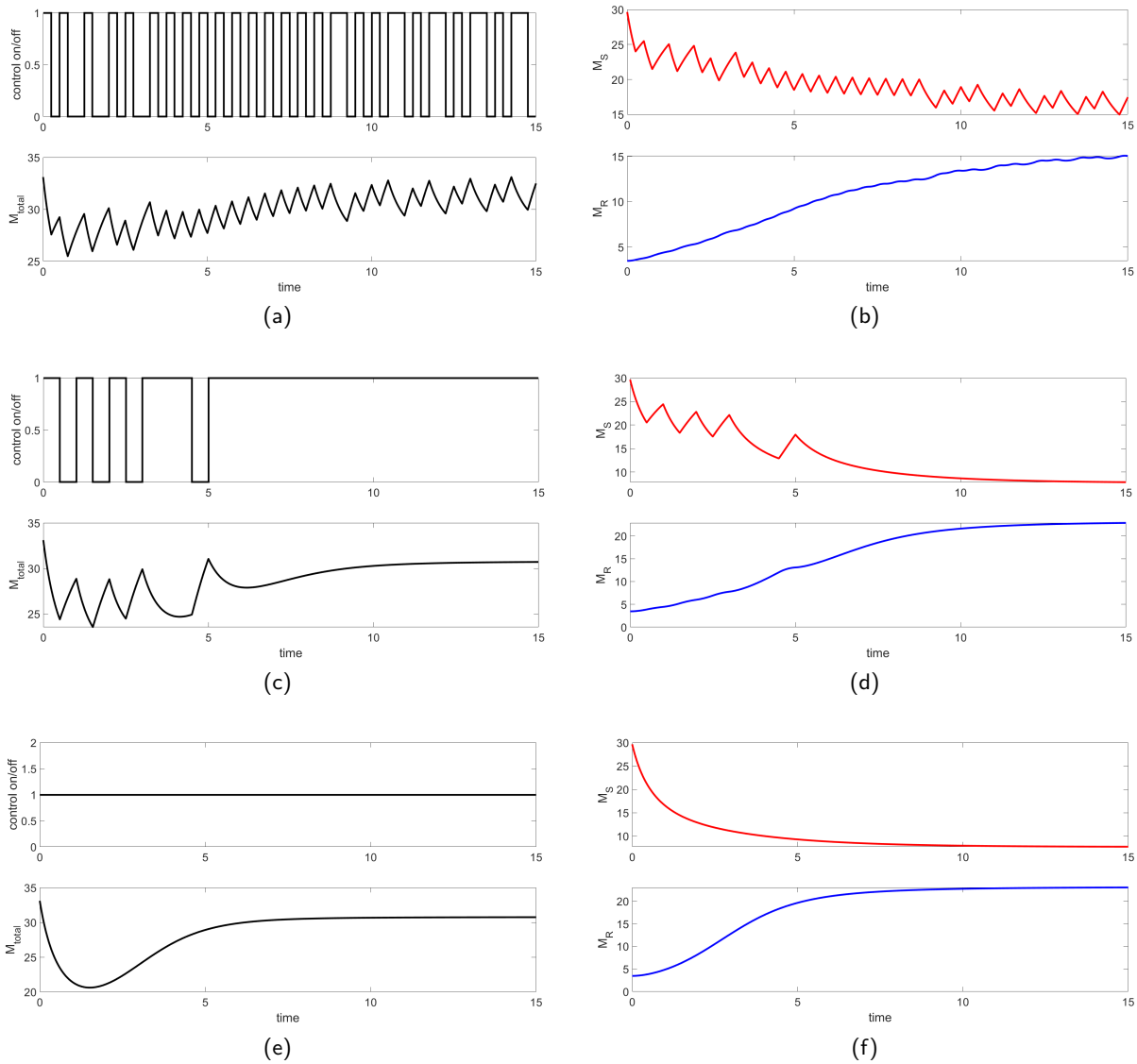


Figure 7-1: Integer control on the Haploid mean field model determined on each sub-interval sequentially, for sub-intervals of length (a)/(b) 0.25 (c)/(d) 0.5 (e)/(f) 1. Control cost determined using parameter set one from table 6.3.

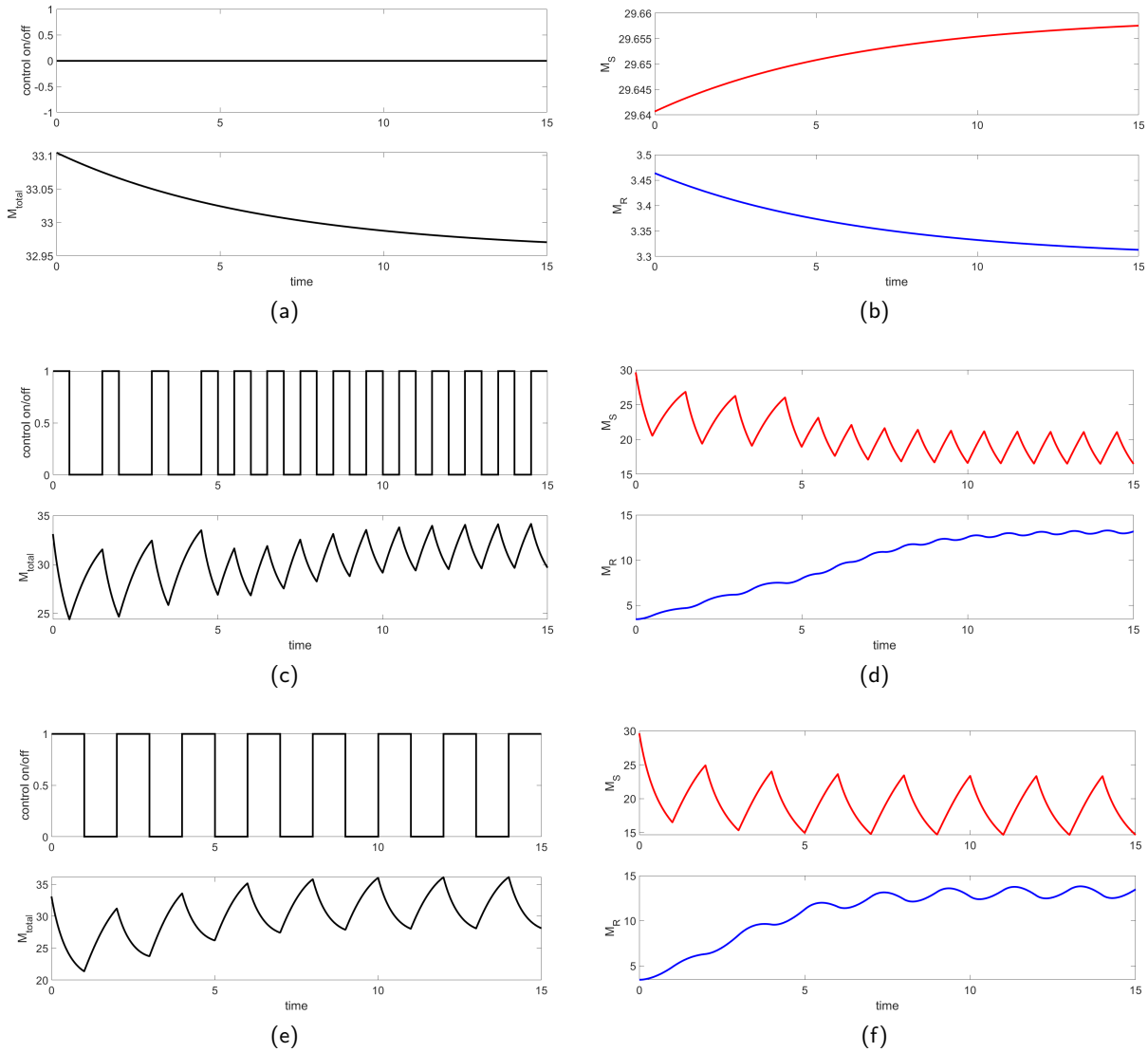
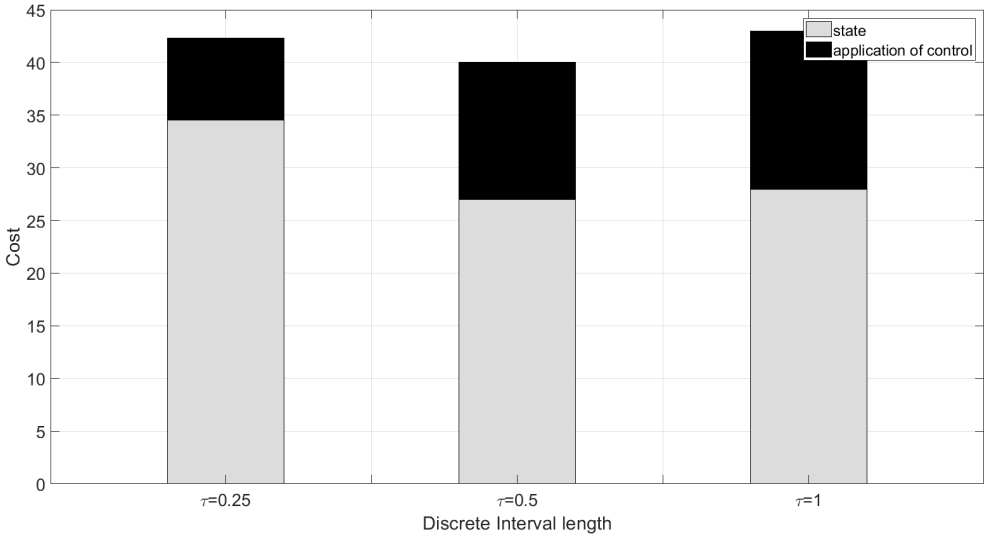
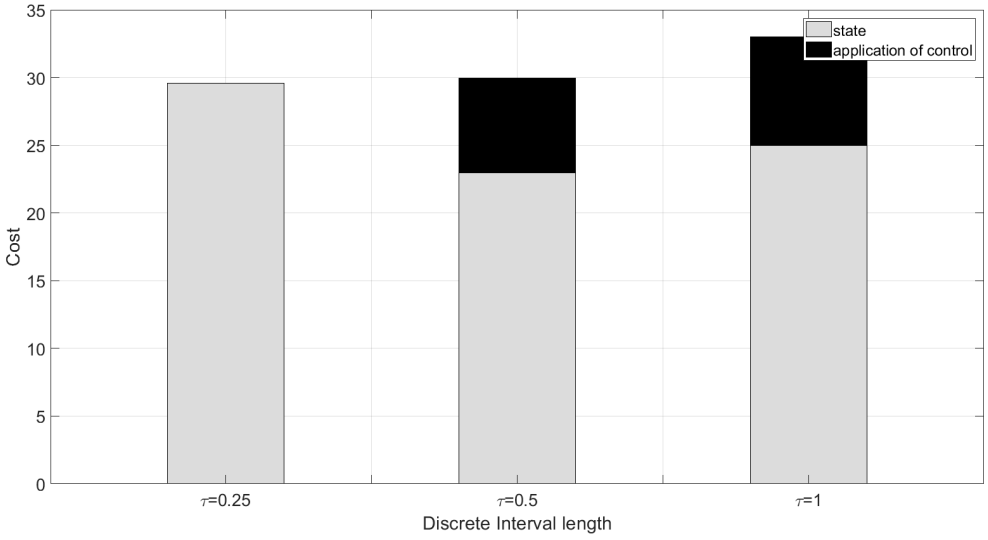


Figure 7-2: Integer control on the Haploid mean field model determined on each sub-interval sequentially, for sub-intervals of length (a)/(b) 0.25 (c)/(d) 0.5 (e)/(f) 1. Control cost determined using parameter set two from table 6.3.



(a)



(b)

Figure 7-3: Cost comparison of the sequential controls on the haploid mean field model for cost function parameter sets one and two for the different length time intervals tested.

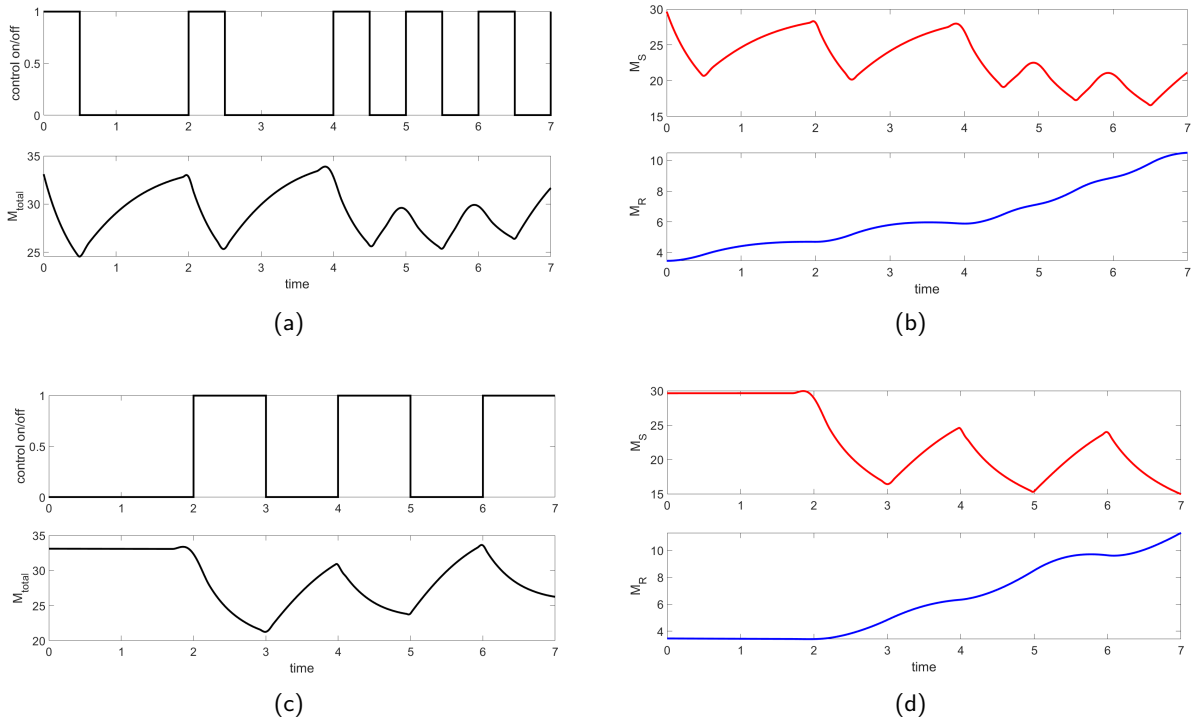


Figure 7-4: Integer control optimised on discrete time intervals of length (a)/(b) 0.5 (c)/(d) 1 for cost function parameters in set two of table 6.3.

again we must limit the length of the overall interval that we use this method on. We limit this interval length to 7, such that any examples will include the behaviour we saw in figure 6-1, with the potential increase in the mean occurring while the hosts are still under treatment. This limits the minimum length of the sub-intervals we may use in the method we use to 0.5.

Figure 7-4 shows examples of these optimal controls for the same cost function as the sequential controls for sub-interval lengths 0.5 and 1. The general trend of the optimised controls shows that the majority of the control applied is applied toward the end of the interval. This can be seen by considering both the length of the controlled periods and the frequency at which they occur. This is the opposite trend to what was seen in chapter 4, where the optimised integer controls typically had control applied early on an stopped in the final few sub-intervals. This result is consistent with the effects seen on the continuous controls for the same models.

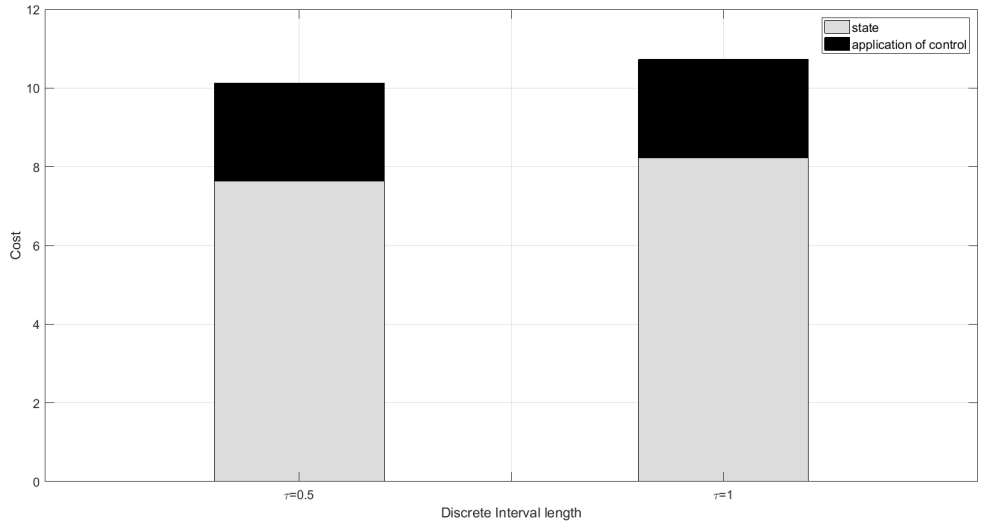


Figure 7-5: Cost comparison of optimal integer controls using different sub-interval length for the haploid mean-field model (6.1). Model parameters given in table 6.1 and cost function parameters given by set two in table 6.3

.

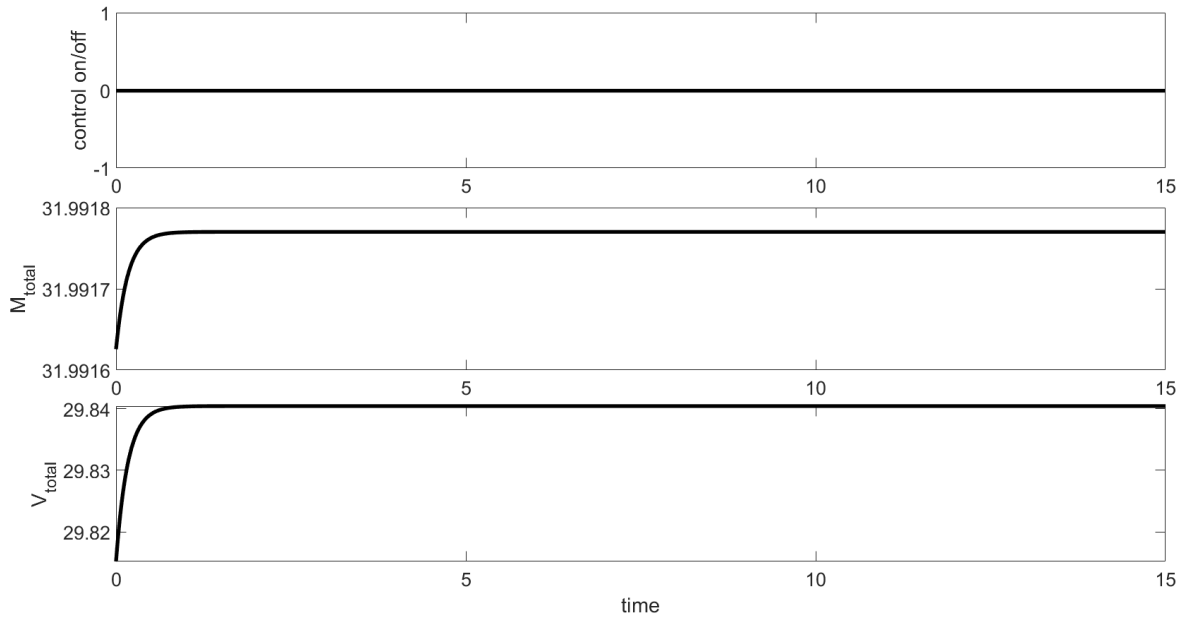
7.2 Haploid Variance Inclusive Model

We perform the same process for the haploid variance inclusive model using the dynamics given in (6.2) with the control limited to $u \in \{0, 1\}$. With the cost function given in (6.7) we may set the model parameters to those given in table 6.1 and determine the sequentially optimised integer control using the cost function parameter sets given in table 6.4. Examples of this for parameter set 4 and sub-intervals of length 0.25 and 0.5 are shown in figures 7-6 and 7-7.

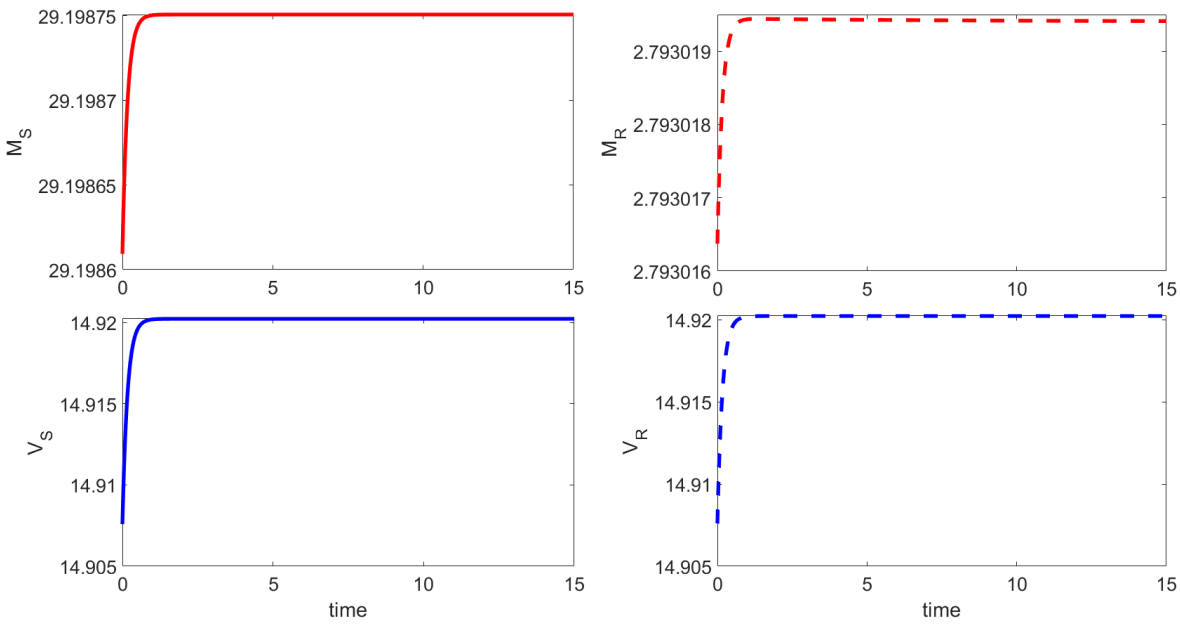
This figure shows that once again the variance inclusive model behaves with some similarity to the mean-field model, with shorter sub-intervals being less likely to have control applied in comparison to longer sub-interval lengths for the same cost function. This is especially clear in this figure where the cost function was minimising the variance. The shortest sub-interval length gives a control that is never applied, while the longest can lead to a control that is applied for the majority of the time. Looking at the comparative cost, figure 7-8, it is clear that treating the population can actually lead to a higher overall cost. This is because the decrease in cost from the lack of susceptible parasites is outweighed by the cost of control itself and the increased cost from the resistant parasite population. The reason why this occurs is exactly the same as we discussed in Section 4.2.1, sequentially determined controls lack the ability to look ahead. Without this, the control can actually lead to increased cost that comes after the interval on which it is applied. This shows that the treatment of macroparasitic diseases should only be done with careful consideration of developing resistance and the future state of infection of the host population. When the proportion of resistance was included in the cost function as well, using the parameters from set 4 in 6.1, control was once again applied even less as expected based on previous results.

Given these results we return once more to examining the discrete time integer treatments optimised across the whole time interval. Figures 7-9, 7-10, 7-11, and 7-12 show examples of these controls for the cost parameters given in sets two and four in table 6.4. Figure 7-13 shows comparisons of the cost breakdowns of each of these controls for the different interval lengths.

The results of this mimic both the mean-field model and the continuous optimal control to an extent, with no control being applied until near the end of the overall interval. Given the response of the model to treatment being stopped, this result is expected as earlier treatment would lead to a greater mean and variance than without treatment and also a greater proportion of resistance in the population, which is what we aimed to minimise.

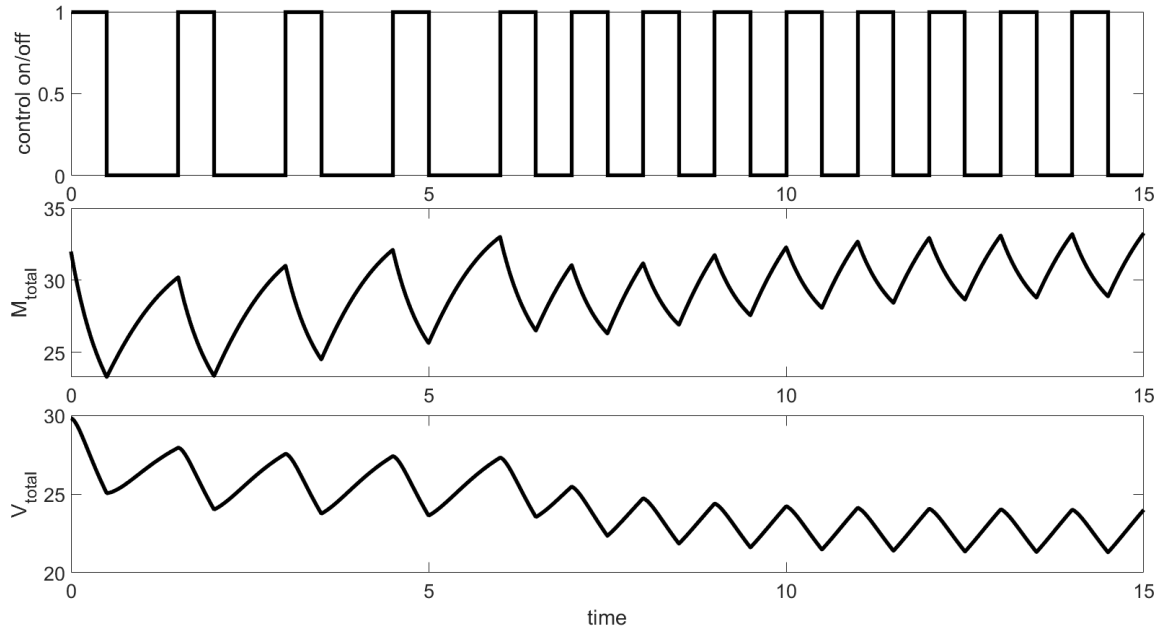


(a)

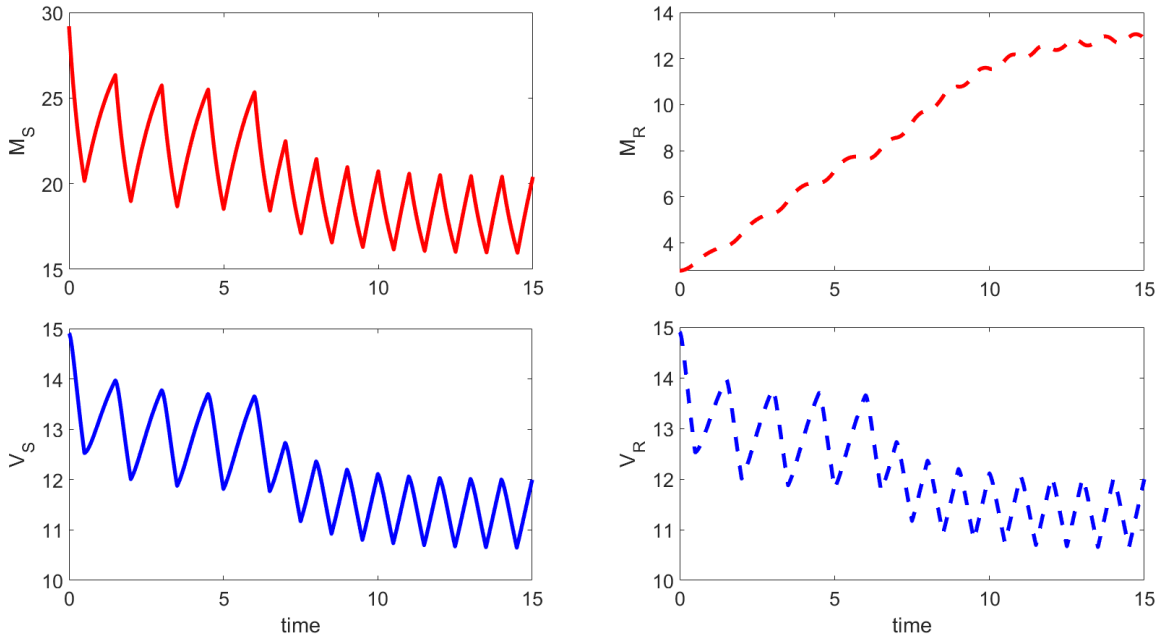


(b)

Figure 7-6: Integer treatment determined sequentially on sub-intervals of length 0.25 for the haploid variance inclusive model (6.2), with model parameters given in table 6.1. Cost function given by (6.7), with parameters given by set 2 in table 6.4.



(a)



(b)

Figure 7-7: Integer treatment determined sequentially on sub-intervals of length 0.5 for the haploid variance inclusive model (6.2), with model parameters given in table 6.1. Cost function given by (6.7), with parameters given by set 2 in table 6.4.

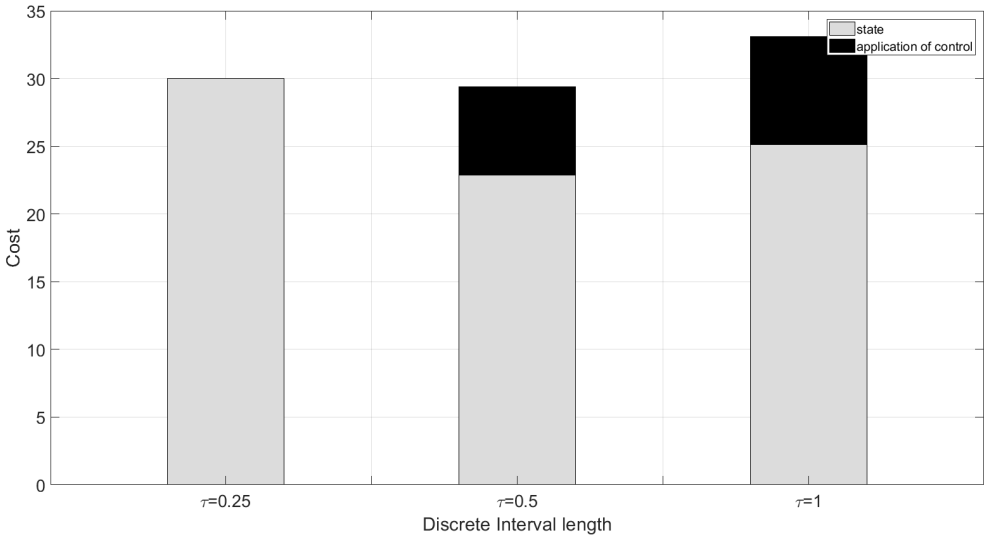
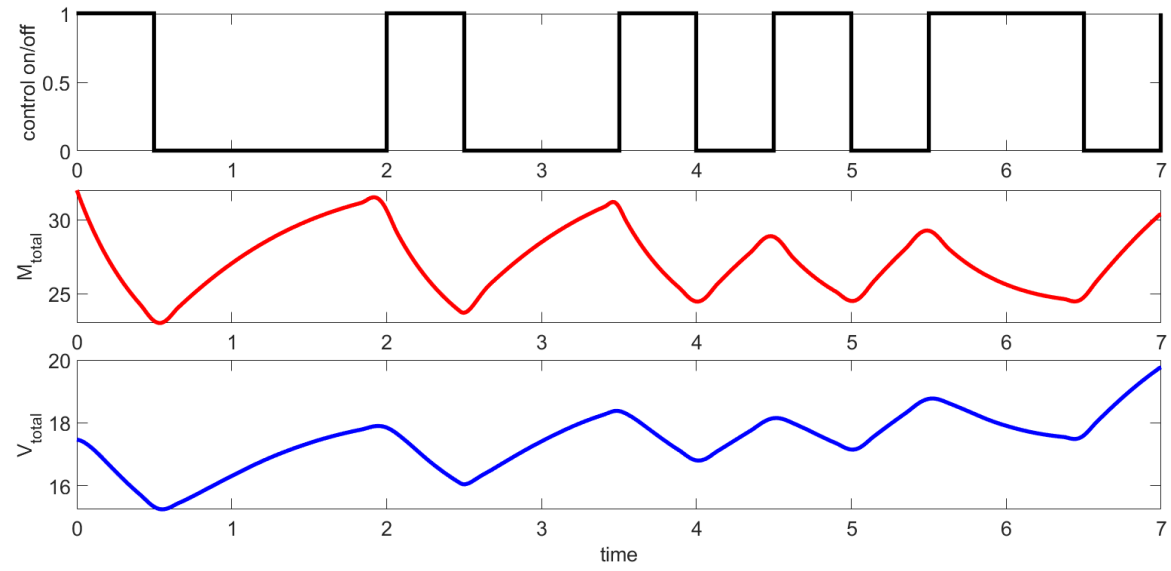
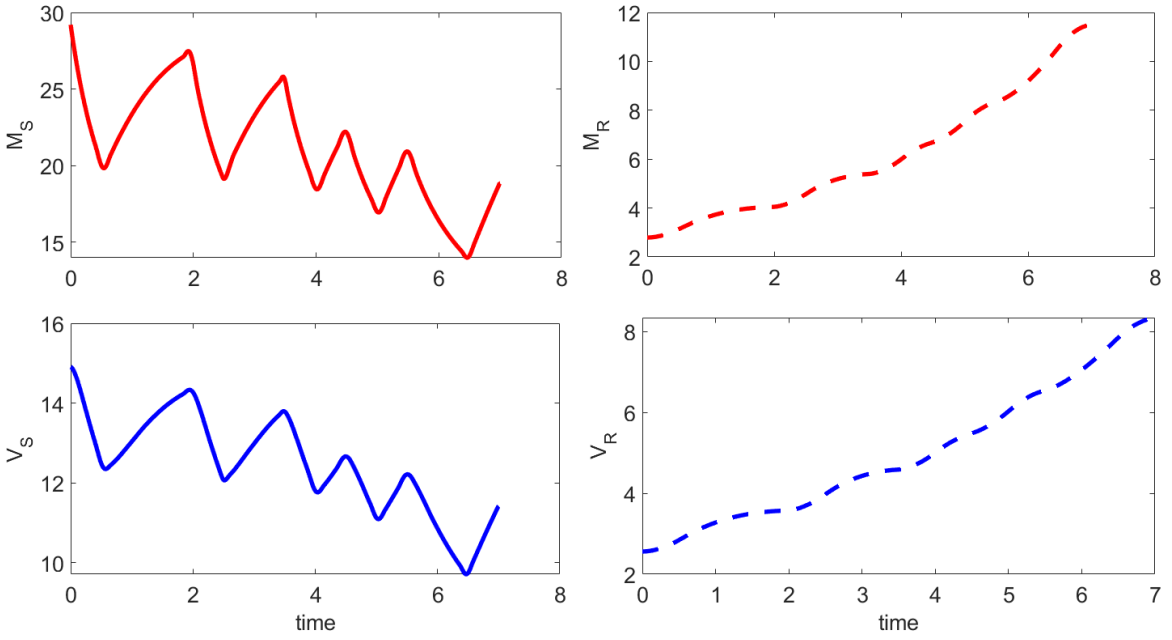


Figure 7-8: Cost comparison of sequential integer controls using different sub-interval length for the haploid variance inclusive model (6.2). Model parameters given in table 6.1 and cost function parameters given by set 2 in table 6.4.

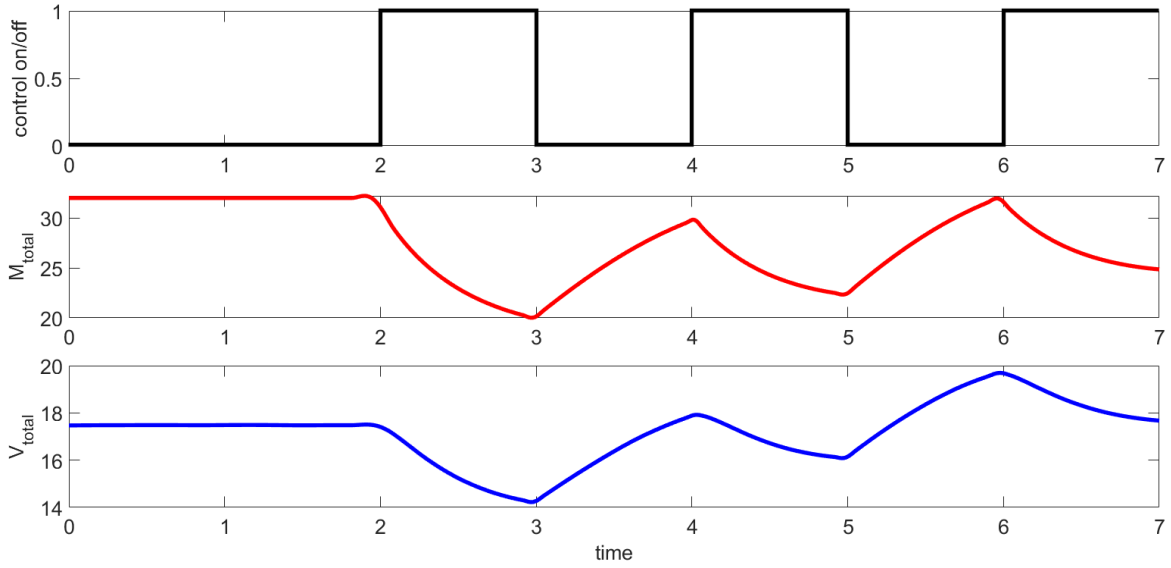


(a)

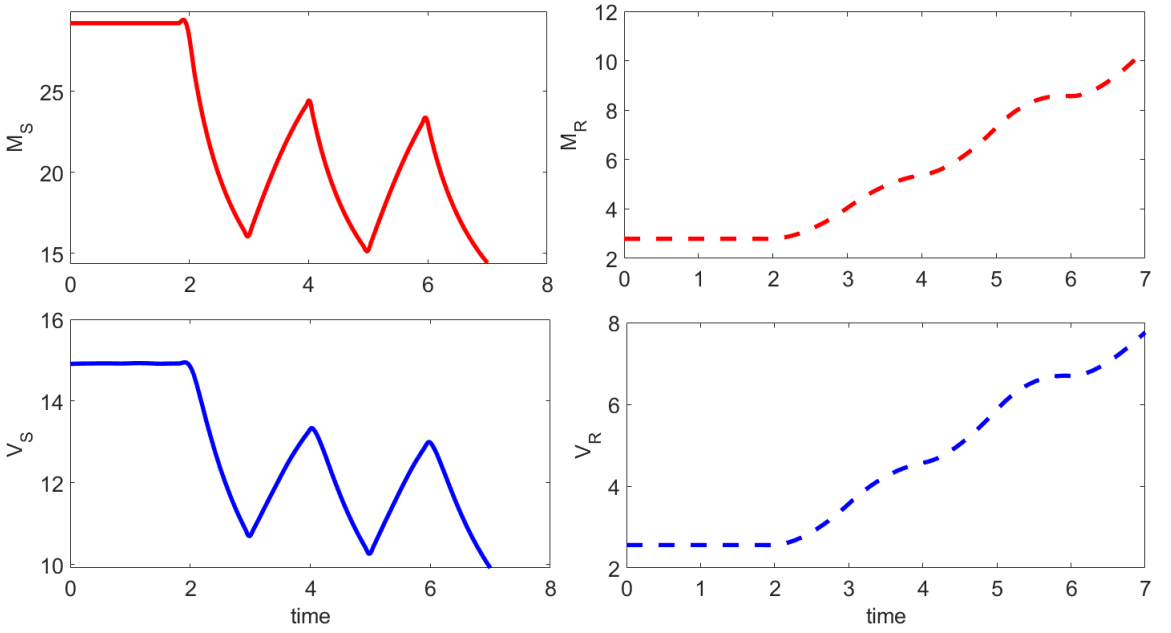


(b)

Figure 7-9: Integer control optimised over full interval on discrete sub-intervals of length 0.5 for the haploid variance inclusive model, with cost function parameters given in set 2 of table 6.4.

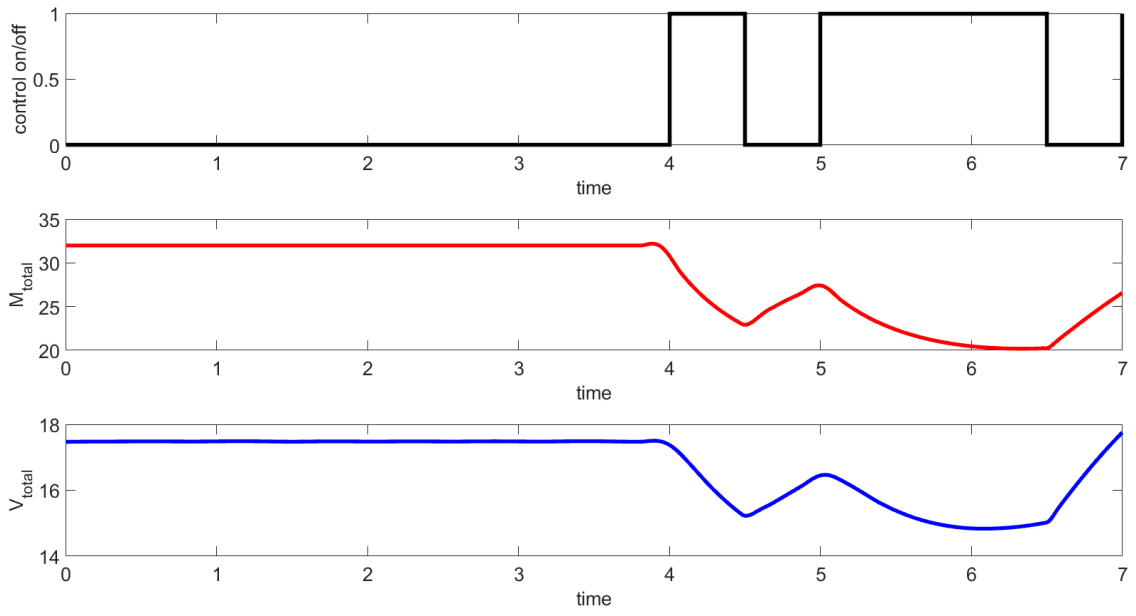


(a)

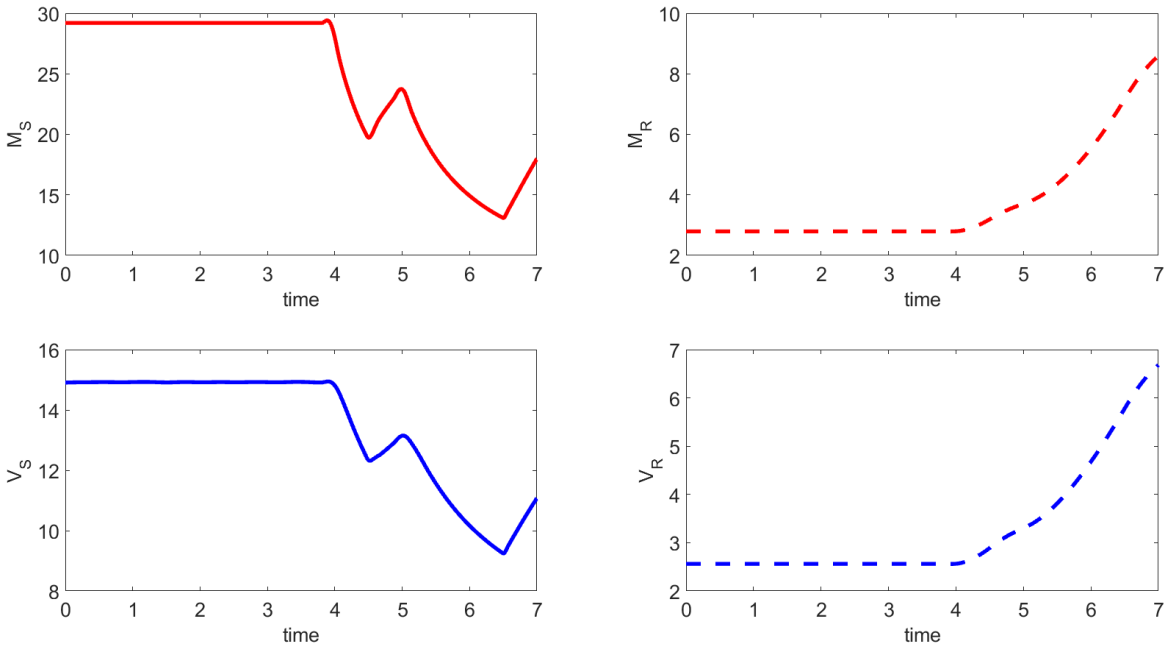


(b)

Figure 7-10: Integer control optimised over full interval on discrete sub-intervals of length 1 for the haploid variance inclusive model, with cost function parameters given in set 2 of table 6.4.

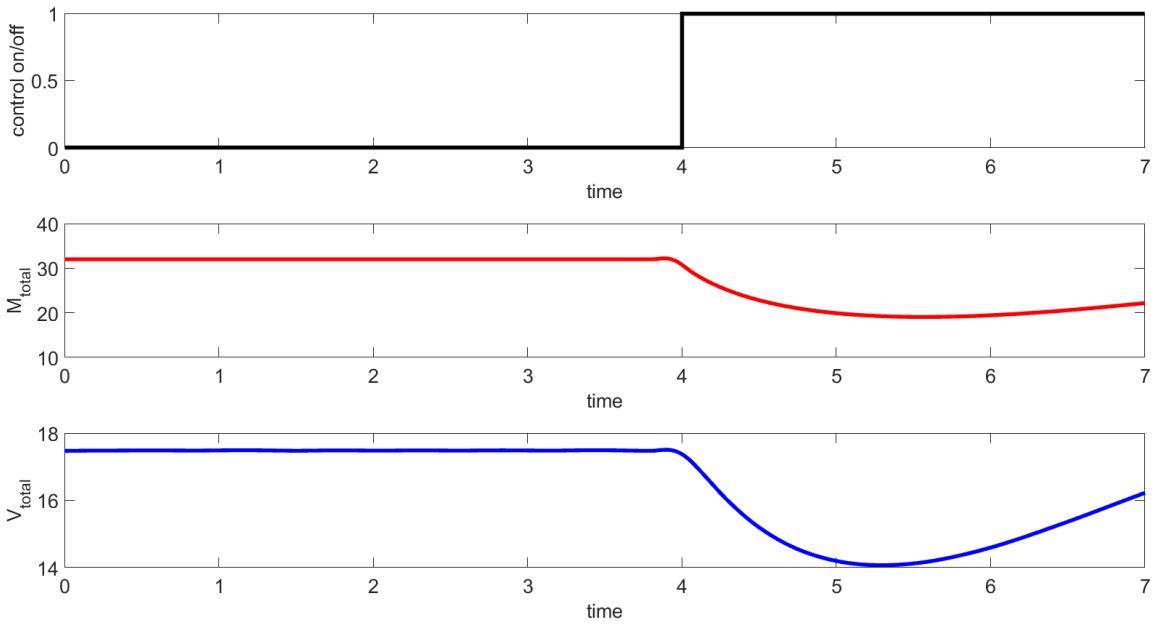


(a)

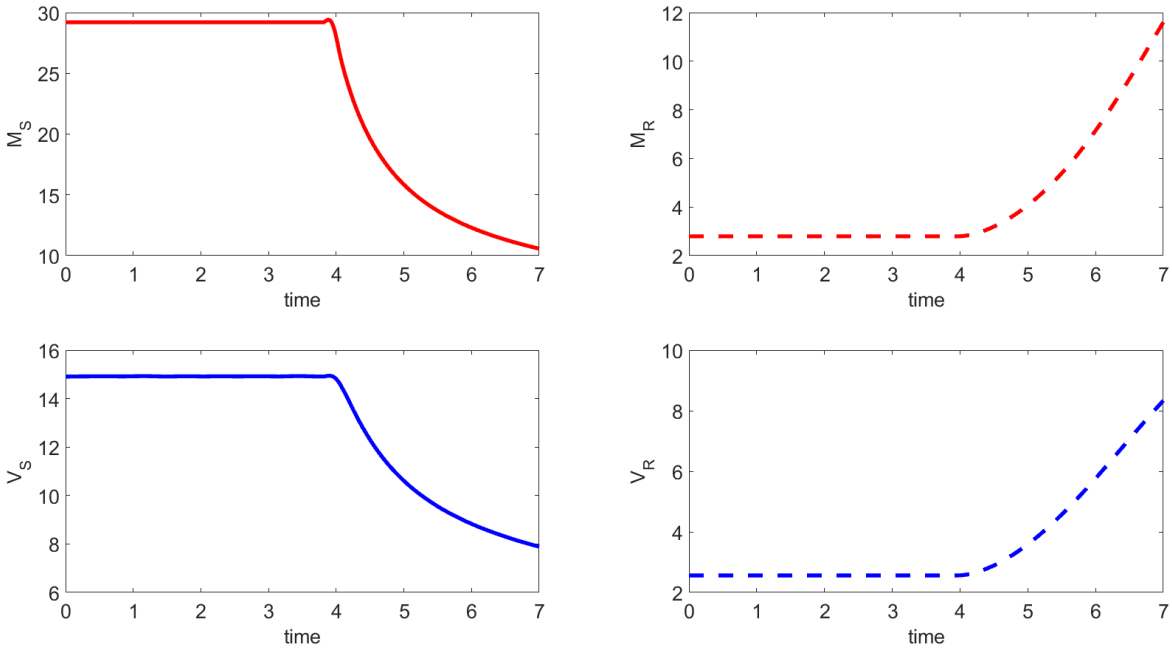


(b)

Figure 7-11: Integer control optimised over full interval on discrete sub-intervals of length 0.5 for the haploid variance inclusive model, with cost function parameters given in set 4 of table 6.4.

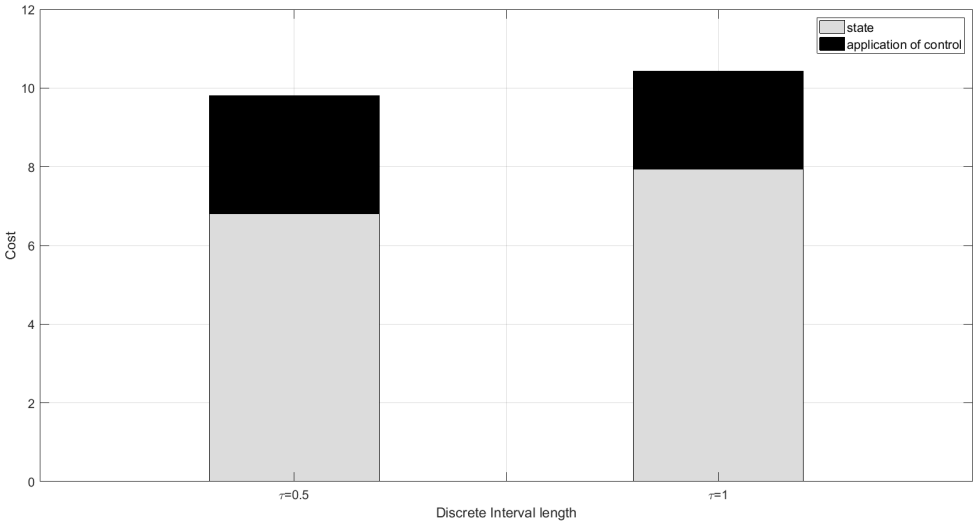


(a)

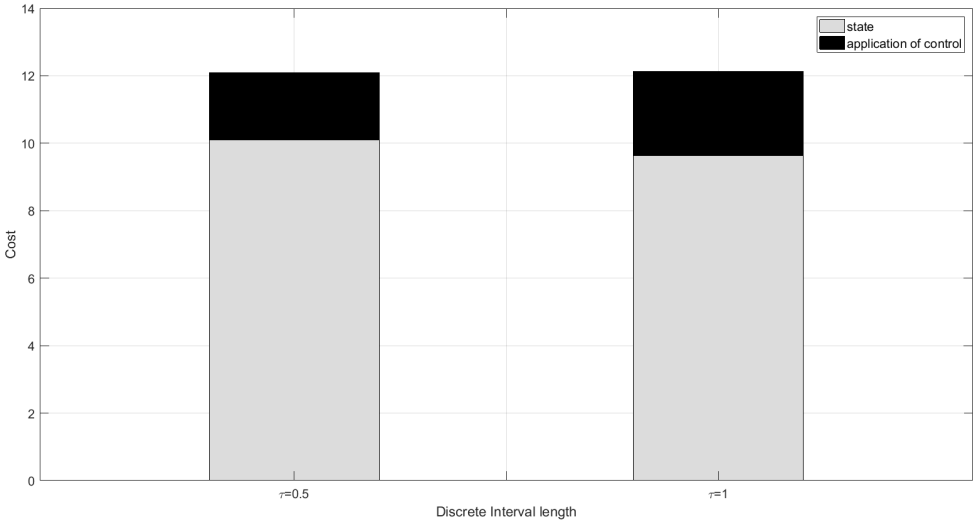


(b)

Figure 7-12: Integer control optimised over full interval on discrete sub-intervals of length 1 for the haploid variance inclusive model, with cost function parameters given in set 4 of table 6.4.



(a)



(b)

Figure 7-13: Cost comparison of optimal integer controls using different sub-interval lengths for the haploid variance inclusive model (6.2). Model parameters given in table 6.1 and cost function parameters given by set (a) two (b) four in table 6.4.

7.3 Diploid Model

Applying integer control to the diploid model is done in the same way as for the previous models. We take the dynamics from (6.3) and the cost function from (6.8). If we set the model parameters to those given in table 6.2, with $\gamma = 1$, then we can determine controls for the situation that the heterozygous parasites are fully susceptible to treatment. If we use the cost function parameters in table 6.5 then by choosing the control sequentially we obtain the examples shown in figure 7-14. Here the figures aim to reduce the mean at the same time as keeping the proportion of resistance in the population low.

If we assume that the heterozygous parasites are resistant to treatment and $\gamma = 0$, then using the cost function parameter sets from table 6.6 for the sequential integer treatments leads to the treatments and dynamics shown in figure 7-15. Once again these sequential controls quickly lead to a dynamic equilibrium between the control and the state with the control being switched on and off in a regular periodic fashion. One of the most important differences between this model and the earlier models is that the length of the sub-intervals was adapted. The examples now look at sub-intervals of length 1.43 and 2 rather than lengths 0.25, 0.5 and 1. This is because much of the behaviour that is of interest when we consider a sustained control happens after the control has been applied for around 7 time units, as can be seen in figure 6-4. We adapted the length of the control intervals such that the optimised integer control will cover a time interval which includes this and will be comparable to the sequential controls.

Examining figures 7-16 and 7-17 we see the optimised integer controls for two cases, one where $\gamma = 1$ (figure 7-16) and another where $\gamma = 0$ (figure 7-17). These controls show far more switching in the control than in the basic mean-field model or the haploid mean-field model. This is likely due to the term in the cost function which penalises the build up of resistant alleles in the population. We saw that the optimal control of the diploid model favoured a greater level of treatment throughout the interval with a drop off at the end. So it would be more expected to see a similar pattern here, however, due to the additional term in the cost function the treatment is switched off instead. This hypothesis is backed up by the results that we saw when using the parameter sets for the cost functions which set $\alpha_r = 0$. These cases are shown in figures G-1 and G-2 in an appendix so that they may be compared. When the proportion of resistant alleles is not penalised, control was applied across the majority of the treatment period and switched off for the end.

It is noted that even in figures 7-16 and 7-17 that when the control is switched on and off throughout the interval that the total mean still never exceeds its value at the untreated equilibrium. In comparison, the haploid models often saw a build up of parasites when treatment was continuously stopped and started. This is as a result of the same property that affected when the control was most intensely applied in the optimal controls, the reactivity of the summation of the variables.

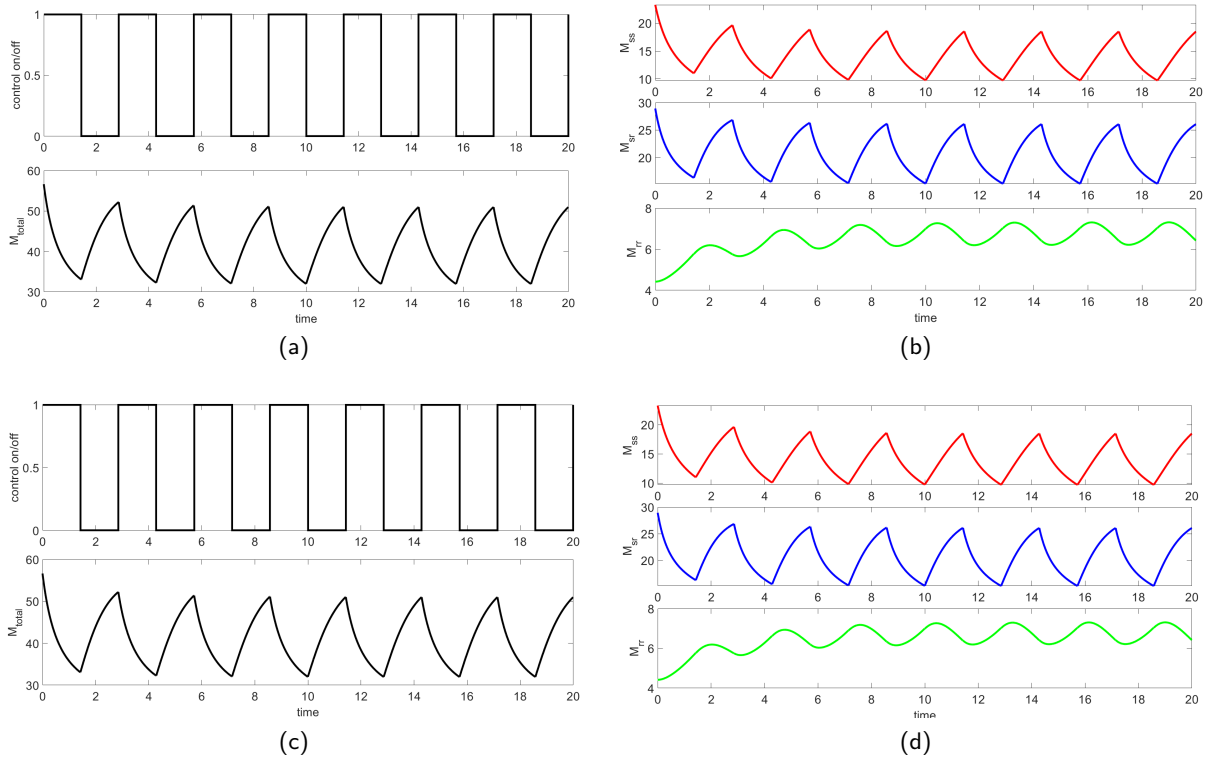


Figure 7-14: Sequentially determined control on the diploid model with model parameters given in table 6.2 with $\gamma = 1$ and cost function parameters in set two of table 6.5. Each sub-interval has a length of (a) $\frac{20}{14}$ (b) 2.

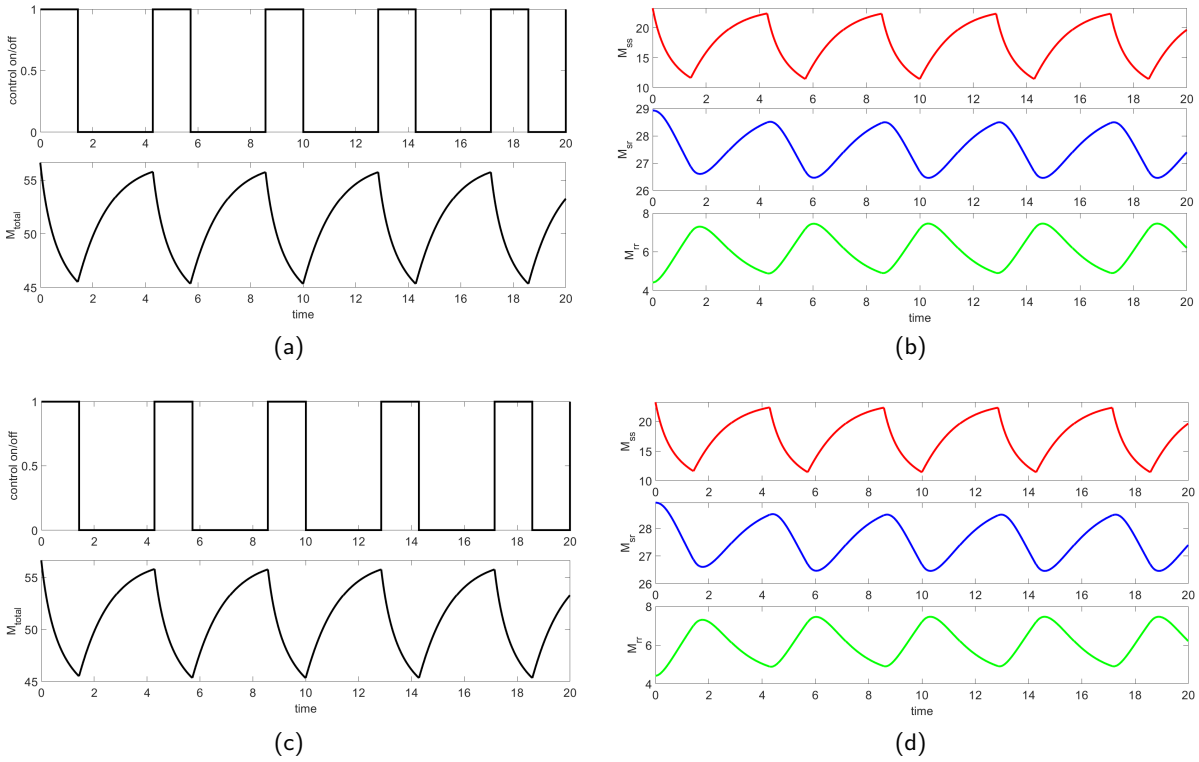


Figure 7-15: Sequentially determined control on the diploid model with model parameters given in table 6.2 with $\gamma = 0$ and cost function parameters in set 4 of table 6.6. Each sub-interval has a length of (a) $\frac{20}{14}$ (b) 2.

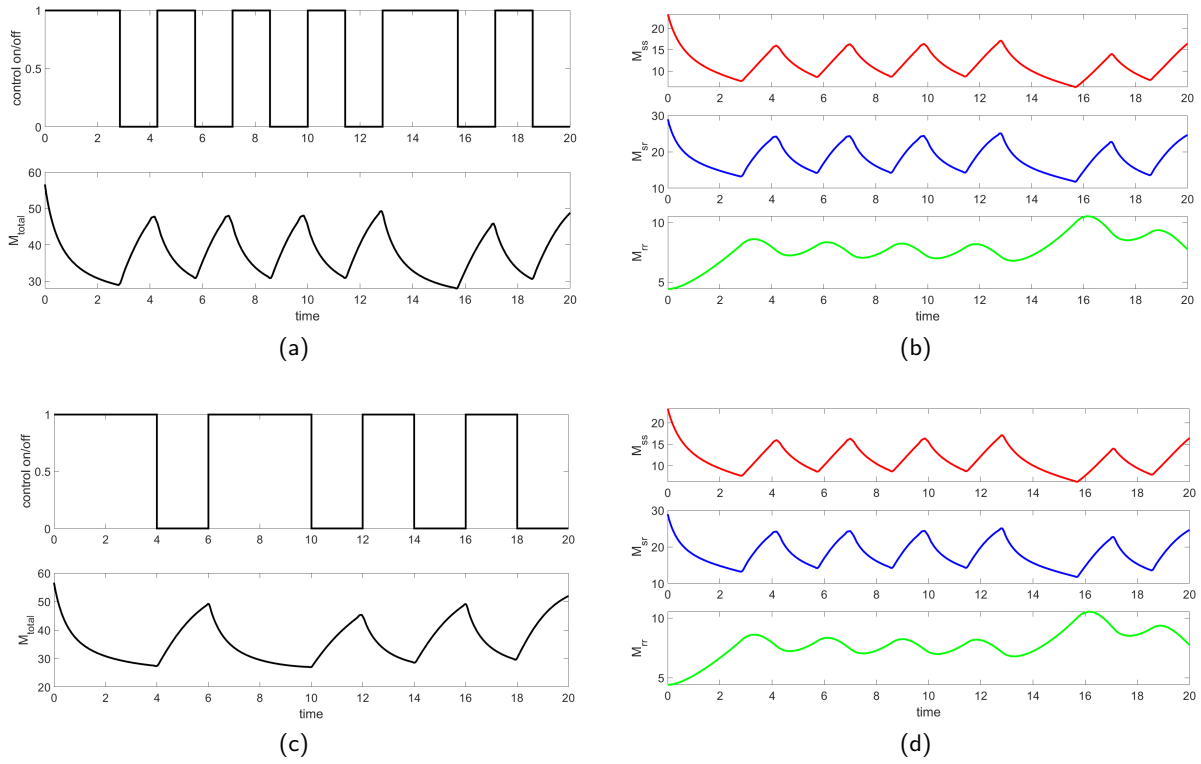


Figure 7-16: Integer control on sub-intervals of length (a)/(b) $\frac{20}{14}$ (c)/(d) 2 optimised across the full interval for the diploid mean-field model with heterozygous parasites acting as resistant, $\gamma = 1$. Cost function parameters used for optimisation given by set 2 in table 6.5.

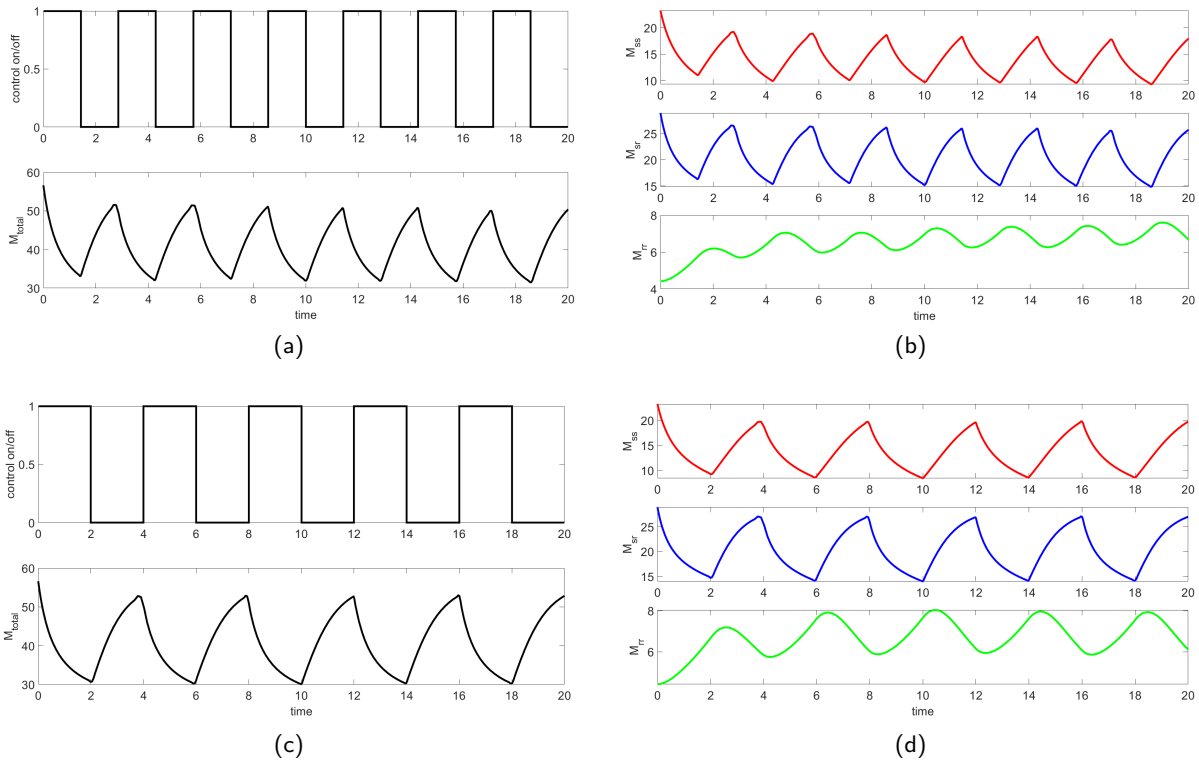


Figure 7-17: Integer control on sub-intervals of length (a)/(b) $\frac{20}{14}$ (c)/(d) 2 optimised across the full interval for the diploid mean-field model with heterozygous parasites acting as resistant, $\gamma = 0$. Cost function parameters used for optimisation given by set 4 in table 6.6.

7.4 Discussion

This chapter serves to extend the study of treatment on the resistant models to consider integer controls in order to examine more realistic parasite treatments. In doing so this chapter has shown that for the haploid models the discrete time integer controls follow a similar, if more exaggerated, trend that the continuous controls do. For the diploid models, the optimised integer controls are far more greatly affected by whether the proportion of resistance is being directly measured and minimised in the cost function. When it is the control is switched on and off to balance between the two aims of the cost function.

The sequentially determined controls show a greater demonstration of the model properties that result in these controls. The haploid models show that, following control, the total variables can reach a state that is greater than if they were untreated, whereas the diploid model returns more directly to equilibrium following treatment. The haploid model controls, in particular, demonstrate the danger of determining control based solely on the present infection level without adequate time for the resistant parasites to die out again.

If we compare the optimised integer controls for these models to those seen for the single genotype model (chapter 4) we see a very similar trend to what was observed for the continuous optimal controls. The haploid models, which show a greater amount of reactivity in the total mean and variance favour control at the end of the overall interval. The diploid model however is where this differs. The optimal controls show far more periodic switching of the control. This occurs due to a combination of the lack of reactivity on the total mean and the need to minimise the proportion of resistance that is in the population. As the comparison with the figures in the appendix show, when the proportion of resistance is not measured in the cost function, the trend is again similar to the model's optimal controls. This means control is applied with the greatest intensity at the start, and continued throughout the majority of the interval, and switched off at the end.

Considering these results, it is clear the impact of resistance on the controls depends primarily on two factors; how the reactivity of the total variables affects the dynamics when the parasite population is perturbed away from equilibrium, and the proportion of resistance which builds up in the population. The single genotype does not have reactivity of the total variables or any build-up of resistance. As such the controls that we saw in chapter 4 favoured a sustained control, as these models were not concerned that resistance would build up and there were no negative effects, aside from the cost, on the state as a result of control being applied. Comparing the haploid and diploid models we found that, although treating the diploid model still led to a build-up of resistance, in general, the diploid model was far less reactive in terms of the total mean and variance than the haploid models. While this may be partially due to parameters, we theorise that it is because the presence of each of the genotypes does not just affect the mortality of the other genotypes but also the birth rates. This levels out the effect of changing the population of one genotype on the others and prevents the model from being as reactive. This lack of reactivity meant that the only

negative that resulted from applying control was the increase in the proportion of resistance and not an increased mean. The optimal control was then given by applying control periodically. This meant that there were periods where the state would be reduced as control was applied but when the cost of resistance outweighed this the control was stopped to allow this to decrease. Finally, the haploid models suffered from an increase in the proportion of resistance as control was applied but were also very reactive in terms of the total variables. As a result, when control was applied it would be initially beneficial, but eventually, resistance would build up and the parasite population with it. If the control was stopped due to this resistance, the reactivity meant that the total states would increase beyond their untreated levels having an even greater negative effect. As such the controls showed no control until late enough in the overall interval that control could be applied without the cost of resistance outweighing the benefit of the minimised state prior to the end of the interval.

The results seen in this chapter reflect previous results seen in studies of drug resistance response to treatment, or the absence thereof [103, 60]. The reduction of resistance when treatment is stopped will be dependent on the competition between non-resistant and resistant parasites and the time scales involved and that for a reduction to be seen the stop treatment must be sustained for long enough to enable the competition between genotypes to have sufficient effect.

Chapter 8

Conclusions and Discussion

Having conducted this research we have considered numerous models and explored how the differences in these models may affect the optimal control. Some of the key elements of this research have been as follows:

- In chapter 2 we examined a previous model and the effects of different control strategies on the stochastic model.
- Chapter 3 included a formal description of a method by which a variance inclusive model could be formulated for linked population models.
- Following this chapter 3 presented an analysis of the models, examining how including the variance can alter the equilibrium values and model response to perturbations in both the long and short term.
- Using the models from chapter 3 we determined controls on them, including continuous optimal controls and integer controls to examine how these model differences affect basic control strategies.
- Chapter 5 considered how to further the model formulation so as to include resistance and competition. This included analysis of the models and comparison between them to see how the parasite resistance to treatment may alter the model.
- Chapters 6 and 7 used the resistance inclusive models and applied control to the dynamics, using the same methodologies as for the original models. This was done to see how having resistance build up may affect the treatment strategies that we saw on our basic models.

8.1 Key Results

Here we review the key results that we found in the course of this research and discuss what made these finding important and why.

The key results that we found by experimenting with different treatment strategies on the stochastic model of Isham was that using individualised treatments could lead to disproportionate changes in the variance when compared with the change in the mean. However, it was also noted that these individualised treatments could not always be performed on the ODE approximation of the simulation model. Nonetheless, it was this finding which drove the research that followed into whether considering the variance as a priority in determining optimal treatments would alter the resultant strategies.

In the formulation of our models in chapter 3 we used a methodology first described by Keeling [52]. Conducting this research did not lead to all positive results and was not without setbacks, chief among which was the nature of the model that we formulated using this method. We found that in spite of using a moment closure based on the relationship between moments seen in a negative binomial distribution this is not enough to force the model dynamics to inherit the properties of this distribution, such as $V > \overline{M}$. Through comparison with the model developed by Isham, we theorise that in order to formulate a model that can fully include these properties further information is needed in regards to host heterogeneity in rate parameters. This is a challenging addition to make to the model as it either requires determination of rate parameters for every host, or for distributions of the parameters to be determined. Both of these involve considering the individual host burdens. Determining individual rate parameters for hosts is a large part of the reason that either fully stochastic or classic two population style metapopulation models may not be used, and using parameter distributions also requires a study of the covariance between the parameters and the individual host burdens.

In spite of this, we used the model we derived and the Isham model to study how controls based on the minimisation of the variance may differ when compared with controls that are intended to minimise the mean. This allowed us to study the differences in control that would arise when the control was based on the mean or the variance on models that predicted both over and under-dispersion. As far as treatments on the whole host population we found that treatments intended to lower the mean and the variance in parasite burden did not qualitatively differ from one another. In the models studied the behaviour of the mean and variance under treatment were very similar to one another in the long term, that is they dropped to lower equilibrium values. The cost function for each of these controls aimed to minimise either the mean, the variance, or a combination of the two, toward a lower target value. As a result, the desired effect of the control was essentially the same for each variable. This meant that the optimal controls were expected to show similar responses, with any difference between them mainly occurring as a result of the specific cost function parameter weightings chosen.

Perhaps one of the most important features of the control that was demonstrated using these

models was the meaning of the control variable. As was shown on the mean-field model and our variance inclusive model, using a control variable to represent a proportion of the host population being treated, as is commonly done for microparasitic disease models, introduced discontinuities and made optimisation difficult. The alternative control variable, which represented the control intensity applied to all hosts, affected the problem in two ways. Firstly, it meant that the model could not be used for individualised treatments, it was only used for treatments across the whole host population. Secondly, it altered what may be considered an admissible control. While the initial optimisations assumed the control variable could take any value within the bounds of $[0, 1]$, we considered the scenario that the control intensity would only be able to switch between two values, corresponding to times when treatment is applied and when it is not. This arose as the intensity of controls is set by the dosage, which is typically fixed to ensure maximum efficacy. By studying this we were able to show that optimising the discrete time controls in this scenario showed similarity to the continuous optimal controls, with control applied strongly at the start of the interval before being stopped. The consideration of integer controls also led to important results in regards to the resistant models.

Focusing on the other side of our investigation we took the methodology we described in Chapter 3 and adapted it to allow for the inclusion of the genetic resistance to treatment. In doing so we exposed potential pitfalls of the formulation method, namely that rate functions that involved reciprocals of our individual parasite populations were not compatible with models that directly modelled the variance. Despite this we were able to use the methodology, combined with simplifying assumptions, to formulate models which assumed that resistance to treatment was determined by a genetic trait for parasites that had both haploid and diploid genetics. Using these models to examine control led to some of the more interesting results. Most importantly was that even in models which are not reactive when the individual variables are considered if the sums of variables, M_{total} and V_{total}, were considered they showed traits such as transient growth, as would occur in a reactive model. The implications of this on control were numerous. Most obviously, treatments could be more effective when applied for shorter periods as initial reductions were greater than the long term reductions. There were also negative implications. The haploid model in particular, showed that periodically starting and stopping treatment, as was done for some integer controls, could lead to reduced efficacy of the treatment and phases of greater infection levels than a completely untreated host population would see.

The combination of these properties gave interesting results when the optimal control was studied. The haploid model showed that, if total variables displayed a large amount of transient growth after treatment was stopped, it was best to delay treatment at higher intensities until near the end of the treatment interval. In contrast, the diploid model showed far less transient growth after treatment was stopped and so the control was optimised by earlier treatment. It was also observed that it was more common for the optimal controls, in the discrete time integer case, to switch on and off on the resistance models. We relate this back to the transient growth of the total means and variances when we apply treatment as it was more beneficial to apply treatment for a short period and stop when or before it started to have a negative effect. This meant that treatment could be

applied when it was having the desired effect, but when it began to have negative consequences it was stopped to avoid the additional cost.

In conclusion our key results suggest that in the models that we formulated the inclusion of the variance had little impact and did not provide sufficient information to more accurately estimate how treatments affect the distribution of parasites. The benefit of this finding is that the variability of the parasite burdens in individual hosts is of less concern when trying to determine controls on the whole host population. In regards to the impact of resistance on treatments, the most important results showed that the way in which parasite resistance to treatment is modelled and the effects that it has on the total parasite populations could have significant effects on the optimal treatments.

8.2 Future Work

In the course of this research we identified multiple avenues which may prove interesting for further work. Some of these hinge upon gaining a greater understanding of the driving processes of parasitic infections and the individualised nature of these with respect to the hosts.

The first of these is in regards to further understanding the effects of heterogeneity in rate parameters between parasite sub-populations. This would be comparable to the work undertaken by Isham and Herbert, examining in greater detail how altering rate parameters can change the distribution of parasites across host populations with a view to determining how the traits of parameter distributions affect the mean and variance. This could in turn lead to models which may better model the distributions of linked populations when there is already empirical evidence to inform the model. The difficulty in doing this lies within the complexity of multiple parameters of different distributions and the lack of data regarding singular processes of parasite population dynamics.

The second of the ideas for further research is into the treatment of only a subset of hosts. While the models we formulated were not suitable for individualised treatments, this research would be intended to build on the work of Isham, and, through a combination of stochastic simulation models and deterministic approximation models, examine in greater detail whether a deterministic model could be used to accurately model the scenario where only some hosts are treated. This is a course of research that we have begun to preliminarily explore using the Isham model. This preliminary research is discussed in the following section.

8.2.1 Split Control on the Isham model

As we discussed for the mean-field and variance inclusive model the interpretation of the control parameter is very important in these models. For all the models presented previously our control parameter, u, has been an increase in the linear death rates of the parasites, which results from a control applied to all hosts simultaneously at a variable intensity. In this section, we discuss

the potential for the control parameter to be a proportion of hosts in the overall population who are subjected to a treatment of fixed intensity. Using the ideas discussed in section 4.1.4 we may formulate a model for the treated and untreated hosts. This is given by:

$$\begin{aligned}
\frac{d\overline{M}_T(t)}{dt} &= \phi h'(1) - \alpha V_T(t) - (\mu_M + \mu_T)\overline{M}_T(t) \\
\frac{dV_T(t)}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M + \mu_T)\overline{M}_T(t) - 2(\mu_M + \mu_T)V_T(t) \\
\frac{d\overline{M}_U(t)}{dt} &= \phi h'(1) - \alpha V_U(t) - (\mu_M)\overline{M}_U(t) \\
\frac{dV_U(t)}{dt} &= \phi(h''(1) + h'(1)) + (\mu_M)\overline{M}_U(t) - 2(\mu_M)V_U(t).
\end{aligned} \tag{8.1}$$

As the hosts in the Isham model are entirely independent of one another the proportion of hosts being treated is not contained in the system itself but the treatment intensity applied to the hosts being treated, μ_T, is. Where the problem arose in the variance inclusive model was in determining the discontinuity that would occur when a host went from untreated to treated or vice versa. For this model, we propose using the mean and the variance to estimate the parameters of a negative binomial distribution. Before this is done we make some simplifying rules for the use of the method, these are

- Hosts are only moved from untreated to treated or vice versa at discrete time points.
- Hosts are not chosen randomly or singularly, they are chosen according to some criteria, for example, we treat those infected above a set level m.
- At the end of treatment, all hosts are moved back to the untreated category, prior to any hosts being selected for the next treatment.

With these set out the method by which the discontinuities are calculated using the following methods: **Moving treated hosts to untreated category**

- Set proportion of hosts undergoing treatment to $p_T = \frac{N-k}{N}$, where N is the number of hosts.
- The total mean becomes
$$\overline{M}_{tot} = p_T\overline{M}_T + (1 - p_T)\overline{M}_U. \tag{8.2}$$
- The total variance is calculated by considering the equation $V_{tot} = \frac{1}{N}\sum_{i=1}^{N}([M_i] - \overline{M}_{tot})^2$ and indexing the treated hosts by the $1 \leq i \leq N-k$ and the treated hosts using $N-k+1 \leq i \leq N$.

Taking

$$
\begin{aligned}
V_{tot} &= \frac{1}{N}\sum_{i=1}^{N}([M_i]-\overline{M}_{tot})^2 \\
&= \frac{1}{N}\left(\sum_{i=1}^{N-k}([M_i]-p_T\overline{M}_T+(1-p_T)\overline{M}_U)^2+\sum_{i=N-k+1}^{N}([M_i]-p_T\overline{M}_T+(1-p_T)\overline{M}_U)^2\right) \\
&= \frac{1}{N}\left(\sum_{i=1}^{N-k}[M_i]^2-2p_T(N-k)\overline{M}_T^2-2(1-p_T)(N-k)\overline{M}_U\overline{M}_T+(N-k)p_T^2\overline{M}_T^2\right. \\
&\left.+2(N-k)(1-p_T)p_T\overline{M}_T\overline{M}_U\right) \\
&+\frac{1}{N}\left(\sum_{i=1}N-k+1^N[M_i]^2-2kp_T\overline{M}_T\overline{M}_U-2k(1-p_T)\overline{M}_U^2+kp_T^2\overline{M}_T^2\right. \\
&\left.+k(1-p_T)^2\overline{M}_U^2+2kp_T(1-p_T)\overline{M}_T\overline{M}_U\right)
\end{aligned}
$$

and using $V_T=\frac{1}{N-k}\sum_{i=1}^{N-k}([M_i]-\overline{M}_T)^2$ and $V_U=\frac{1}{k}\sum_{i=N-k+1}^{N}([M_i]-\overline{M}_T)^2$ this may be rearranged to give

$$V_{tot}=p_TV_T+(1-p_T)V_U+(\overline{M}_T^2+\overline{M}_U^2-2\overline{M}_T\overline{M}_U)(p_T(1-p_T)). \tag{8.3}$$

- Set the proportion undergoing treatment to $p_T=0$ and the means and variances to

$$\begin{aligned}\overline{M}_U=\overline{M}_{tot} \quad \overline{M}_T=0\\ V_U=V_{tot} \quad V_T=0.\end{aligned}$$

Moving untreated hosts to treated category

- Start with state conditions

$$\begin{aligned}\overline{M}_U=\overline{M}_{tot} \quad \overline{M}_T=0\\ V_U=V_{tot} \quad V_T=0.\end{aligned}$$

- Calculate distribution parameters from the untreated mean and variance using

$$p=\frac{\overline{M}_U}{V_U} \text{ and } r=\frac{\overline{M}_U^2}{V_U-\overline{M}_U}. \tag{8.4}$$

- Set up the estimated distribution, $NB(p,r)$, using calculated parameters.
- Use the cumulative distribution function of the distribution, in this case

$$\mathbb{P}(M\leq m)=I_p(m+1,r)=\frac{\Gamma(r+m)\int_0^p t^m(1-t)^{r-1}dt}{\Gamma(r)\Gamma(m)},$$

called the regularised incomplete Beta function, to calculate the proportion of hosts estimated

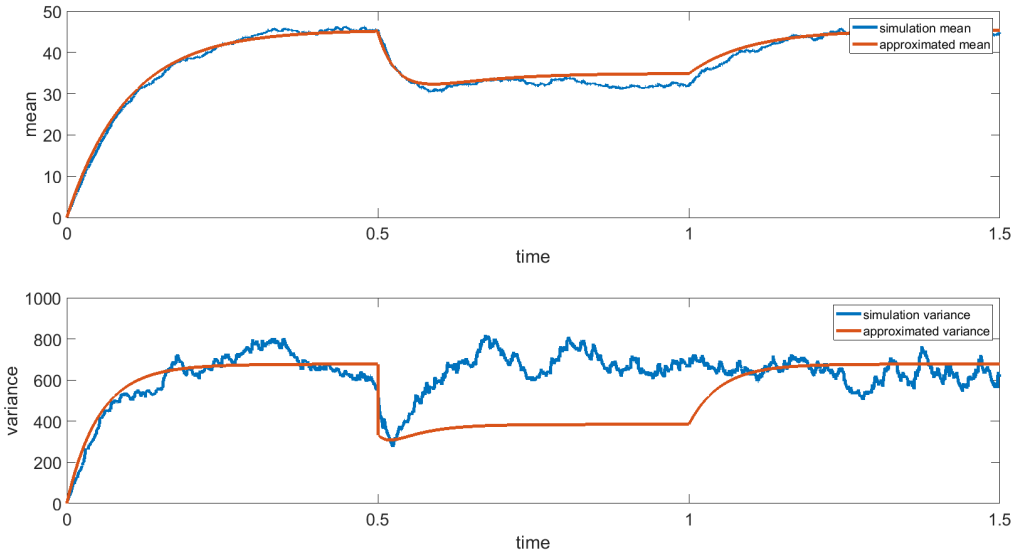


Figure 8-1: Comparison of a timed treatment on the stochastic Isham model on hosts infected with parasite burden greater than 50, at the beginning of the treatment, with a deterministic model using the method described.

to be infected above, p_T, and below, $1 - p_T$, the critical value m.

- Truncate the distribution at the value m and calculate the mean and variance of the lower tail. This gives the new values of $\overline{M}_U$ and V_U.
- Rearrange equations (8.2) and (8.3) to give the new values of $\overline{M}_T$ and V_T.

As a preliminary test, this method is applied for a single timed treatment on the Isham model. This is presented in figure 8-1. This figure shows that the method is initially a good estimate for the stochastic mean. When we look at the variance in the stochastic burden and the deterministic model for the variance it is clear that the deterministic model is a reasonable approximation during the initial treatment. This does not remain the case for long as the results of the stochastic model shows that the variance quickly returns to its original value, even while treatment is applied. This behaviour is not seen in the deterministic model. These initial results show that although the method is by no means finished it has potential and with further exploration could provide useful results.

Bibliography

[1] The b.r.p. cattle and sheep parasite control product guide. http://beefandlamb.ahdb.org.uk/, May 2018.

[2] Antiparasitic resistance. https://www.fda.gov/animal-veterinary/safety-health/antiparasitic-resistance, Feb. 2019.

[3] Adler, F., and Kretzschmar, M. Aggregation and stability in parasite—host models. *Parasitology 104*, 2 (1992), 199–205.

[4] Agriculture, and Board, H. D. The b.r.p. cattle and sheep parasite control product guide. beefandlamb.ahdb.org.uk.

[5] Alexander, N., Moyeed, R., and Stander, J. Spatial modelling of individual-level parasite counts using the negative binomial distribution. *Biostatistics 1*, 4 (2000), 453–463.

[6] Allen, L. J., Brauer, F., Van den Driessche, P., and Wu, J. *Mathematical epidemiology*, vol. 1945. Springer, New York, 2008.

[7] Anderson, R., and May, R. Regulation and stability of host-parasite population interactions: I. regulatory processes. *The Journal of Animal Ecology* (1978), 219–247.

[8] Anderson, R., and May, R. The invasion, persistence and spread of infectious diseases within animal and plant communities. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences 314*, 1167 (1986), 533–570.

[9] Anderson, R., and May, R. *Infectious diseases of humans: dynamics and control*. Oxford university press, Oxford UK, 1991.

[10] Bauch, C., and Rand, D. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. *Proceedings of the Royal Society of London B: Biological Sciences 267*, 1456 (2000), 2019–2027.

[11] Behncke, H. Optimal control of deterministic epidemics. *Optimal control applications and methods 21*, 6 (2000), 269–285.

[12] Bennett, G. F. Studies on cuterebra emasculator fitch 1856 (diptera: Cuterebridae) and a discussion of the status of the genus cephenemyia ltr. 1818. *Canadian Journal of Zoology 33*, 2 (1955), 75–98.

[13] Bennett, J. E., Dolin, R., and Blaser, M. J. *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases: 2-Volume Set*, vol. 1. Elsevier Health Sciences, Amsterdam, 2014.

[14] Bethony, J., Brooker, S., Albonico, M., Geiger, S. M., Loukas, A., Diemert, D., and Hotez, P. J. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *The Lancet 367*, 9521 (2006), 1521–1532.

[15] Boltyanskiy, V., Gamkrelidze, R. V., and Pontryagin, L. Theory of optimal processes. Tech. rep., Joint Publications Research Service, Arlington, VA, 1961.

[16] Bonnans, F., Martinon, P., Giorgi, D., Grelard, V., Maindrault, S., Tissot, O., and Liu, J. *BOCOP 2.1.0 - User Guide*. BOCOP, Oct. 2017.

[17] Brauer, F., Castillo-Chavez, C., and Castillo-Chavez, C. *Mathematical models in population biology and epidemiology*, vol. 40. Springer, New York, 2001.

[18] Churcher, T. S., Fferguson, N. M., and Basáñez, M. Density dependence and overdispersion in the transmission of helminth parasites. *Parasitology 131*, 1 (2005), 121–132.

[19] Colley, D. G., Bustinduy, A. L., Secor, W. E., and King, C. H. Human schistosomiasis. *The Lancet 383*, 9936 (2014), 2253–2264.

[20] Cook, J. D. Notes on the negative binomial distribution. `https://statweb.stanford.edu/~owen/courses/306a-1314/negative_binomial.pdf`, July 2020.

[21] Crompton, D. Accelerating work to overcome the impact of neglected tropical diseases. *A roadmap for implementation. Geneva: WHO* (2012).

[22] Crow, J. F., Kimura, M., et al. *An introduction to population genetics theory.* The Blackburn Press, New Jersey, 1970.

[23] De Silva, N., Chan, M., and Bundy, D. Morbidity and mortality due to ascariasis: re-estimation and sensitivity analysis of global numbers at risk. *Tropical Medicine & International Health 2*, 6 (1997), 513–518.

[24] DeGroot, M. H., and Schervish, M. J. *Probability and statistics*. Pearson Education, London, 2012.

[25] Despommier, D. D., and Karapelou, J. W. *Parasite life cycles*. Springer Science & Business Media, Berlin, 2012.

[26] Diekmann, O., and Kretzschmar, M. Patterns in the effects of infectious diseases on population growth. *Journal of Mathematical Biology 29*, 6 (1991), 539–570.

[27] Douglas, J. Fitting the neyman type a (two parameter) contagious distribution. *Biometrics 11*, 2 (1955), 149–173.

[28] Echevarria, F., Borba, M., Pinheiro, A., Waller, P., and Hansen, J. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: Brazil. *Veterinary Parasitology 62*, 3-4 (1996), 199–206.

[29] Elkins, D., Haswell-Elkins, M., and Anderson, R. The epidemiology and control of intestinal helminths in the pulicat lake region of southern india. i. study design and pre- and post-treatment observations on ascaris lumbricoides infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene 80*, 5 (1986), 774–792.

[30] Enns, R. H., M. G. C. *Computer Algebra Recipes: An Advanced Guide to Scientific Modeling, Phase-Plane Analysis*. Springer, New York, 2007.

[31] Etienne, R. S. A scrutiny of the levins metapopulation model. *Comments® on Theoretical Biology 7*, 4 (2002), 257–281.

[32] Evtushenko, Y. G., and Zhadan, V. G. Stable barrier-projection and barrier-newton methods for linear and nonlinear programming. In *Algorithms for Continuous Optimization*. Springer New York, 1994, pp. 255–285.

[33] Ewens, W. J. *Mathematical population genetics 1: theoretical introduction*, vol. 27. Springer Science & Business Media, Berlin, 2012.

[34] Fairweather, I., and Boray, J. Fasciolicides: efficacy, actions, resistance and its management. *The Veterinary Journal 158*, 2 (1999), 81–112.

[35] Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry 81*, 25 (1977), 2340–2361.

[36] Grenfell, B. Parasitism and the dynamics of ungulate grazing systems. *The American Naturalist 139*, 5 (1992), 907–929.

[37] Grenfell, B., Wilson, K., Isham, V., Boyd, H., and Dietz, K. Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode–ruminant interactions as a case study. *Parasitology 111*, S1 (1995), S135–S151.

[38] Griffiths, R. The application of some anthelmintics in veterinary practice. *Journal of Pharmacy and Pharmacology 6*, 1 (1954), 921–943.

[39] Gryseels, B., Polman, K., Clerinx, J., and Kestens, L. Human schistosomiasis. *The Lancet 368*, 9541 (2006), 1106–1118.

[40] Hanski, I. Metapopulation dynamics: From concepts and observations to predictive models. In *Metapopulation Biology*, I. Hanski and M. E. Gilpin, Eds. Academic Press, San Diego, 1997, pp. 69 – 91.

[41] Hanski, I., et al. *Metapopulation ecology*. Oxford University Press, Oxford UK, 1999.

[42] Hanski, I., and Simberloff, D. The metapopulation approach, its history, conceptual domain, and application to conservation. In *Metapopulation Biology*, I. Hanski and M. E. Gilpin, Eds. Academic Press, San Diego, 1997, pp. 5 – 26.

[43] Harrison, G. Stability under environmental stress: resistance, resilience, persistence, and variability. *The American Naturalist 113*, 5 (1979), 659–669.

[44] Harrison, S. Local extinction in a metapopulation context: an empirical evaluation. *Biological journal of the Linnean Society 42*, 1-2 (1991), 73–88.

[45] Harrison, S., and Taylor, A. D. Empirical evidence for metapopulation dynamics. In *Metapopulation biology*. Elsevier, 1997, pp. 27–42.

[46] Headrick, T. C. *Statistical simulation: Power method polynomials and other transformations*. CRC Press, Florida, 2009.

[47] Herbert, J., and Isham, V. Stochastic host-parasite interaction models. *Journal of mathematical biology 40*, 4 (2000), 343–371.

[48] House, T. Algebraic moment closure for population dynamics on discrete structures. *Bulletin of mathematical biology 77*, 4 (2015), 646–659.

[49] Isham, V. Stochastic models of host-macroparasite interaction. *The Annals of Applied Probability* (1995), 720–740.

[50] Iwasa, Y., and Wada, G. Complex life cycle and body sizes at life-history transitions for macroparasites. *Evolutionary Ecology Research 8*, 8 (2006), 1427–1443.

[51] Junquera, P. Imax cd lv pour-on for beef and dairy cattle. `http://parasitipedia.net/index.php?option=com_content&view=article&id=3758&Itemid=4147`, Jan. 2017.

[52] Keeling, M. J. Metapopulation moments: coupling, stochasticity and persistence. *Journal of Animal Ecology 69*, 5 (2000), 725–736.

[53] Keeling, M. J., and Rohani, P. *Modeling infectious diseases in humans and animals*. Princeton University Press, New Jersey, 2011.

[54] Kirches, C. *Fast numerical methods for mixed-integer nonlinear model-predictive control*. Springer, New York, 2011.

[55] Kocvara, M., and Stingl, M. On the solution of large-scale sdp problems by the modified barrier method using iterative solvers. *Mathematical programming 120*, 1 (2009), 285.

[56] Kretzschmar, M., and Adler, F. Aggregated distributions in models for patchy populations. *Theoretical population biology 43*, 1 (1993), 1–30.

[57] Lee, J., and Leyffer, S. *Mixed integer nonlinear programming*, vol. 154. Springer Science & Business Media, Berlin Germany, 2011.

[58] Lenhart, S., and Workman, J. T. *Optimal control applied to biological models*. CRC Press, Florida, 2007.

[59] Liberzon, D. *Switching in Systems and Control*. Systems & Control: Foundations & Applications,. Birkhäuser, Boston, Boston, MA, 2003.

[60] Mackinnon, M. Drug resistance models for malaria. *Acta tropica 94*, 3 (2005), 207–217.

[61] Marion, G., Renshaw, E., and Gibson, G. Stochastic effects in a model of nematode infection in ruminants. *Mathematical Medicine and Biology: A Journal of the IMA 15*, 2 (1998), 97–116.

[62] Marion, G., Renshaw, E., and Gibson, G. Stochastic modelling of environmental variation for biological populations. *Theoretical population biology 57*, 3 (2000), 197–217.

[63] Mark Viney, J. C. Macroparasite life histories - review. *Current Biology 21* (Sept. 2011), 767 – 774.

[64] Martin, R. Modes of action of anthelmintic drugs. *The Veterinary Journal 154*, 1 (1997), 11–34.

[65] Massé, J.-C., and Theodorescu, R. Neyman type a distribution revisited. *Statistica Neerlandica 59*, 2 (2005), 206–213.

[66] May, R. M. Togetherness among schistosomes: its effects on the dynamics of the infection. *Mathematical Biosciences 35*, 3-4 (1977), 301–343.

[67] May, R. M., and Anderson, R. M. Regulation and stability of host-parasite population interactions: Ii. destabilizing processes. *The Journal of Animal Ecology* (1978), 249–267.

[68] McCallum, H. e. a. Breaking beta: deconstructing the parasite transmission function. *Philosophical Transactions of the Royal Society B: Biological Sciences 372*, 1719 (2017), 20160084.

[69] McGraw-Hill. Concise dictionary of modern medicine, 2002.

[70] Mohideen, M., Perkins, J., and Pistikopoulos, E. Towards an efficient numerical procedure for mixed integer optimal control. *Computers & Chemical Engineering 21* (1997), S457–S462.

[71] Moloo, A. Who data show unprecedented treatment coverage for bilharzia and intestinal worms. `https://www.who.int/neglected_diseases/news/unprecedented-treatment-coverage-bilharzia-intestinal-worms/en/`, Dec. 2018.

[72] Morand, S., Krasnov, B. R., and Poulin, R.and Allen Degen, A.

[73] Morgan, E. R. e. a. Ruminating on complexity: macroparasites of wildlife and livestock. *Trends in ecology & evolution 19*, 4 (2004), 181–188.

[74] Murray, J. *Mathematical Biology: I. An Introduction*. Springer, New York, 2002.

[75] Neubert, M. G., and Caswell, H. Alternatives to resilience for measuring the responses of ecological systems to perturbations. *Ecology 78*, 3 (1997), 653–665.

[76] Nisbet, R. M., and Gurney, W. *Modelling fluctuating populations*. John Wiley and Sons Limited, New Jersey, 1982.

[77] Nolan, T. J., and Smith, G. Time series analysis of the prevalence of endoparasitic infections in cats and dogs presented to a veterinary teaching hospital. *Veterinary parasitology 59*, 2 (1995), 87–96.

[78] Olsen, A. e. a. Strongyloidiasis–the most neglected of the neglected tropical diseases? *Transactions of the Royal Society of Tropical Medicine and Hygiene 103*, 10 (2009), 967–972.

[79] Pennycuick, L. Frequency distributions of parasites in a population of three-spined sticklebacks, gasterosteus aculeatus l., with particular reference to the negative binomial distribution. *Parasitology 63*, 3 (1971), 389–406.

[80] Pinheiro, S. Persistence and existence of stationary measures for a logistic growth model with predation. *Stochastic Models 32*, 4 (2016), 513–538.

[81] Pontryagin, L. S. *Mathematical theory of optimal processes*. Routledge, Abingdon UK, 2018.

[82] Prichard, R. Anthelmintic resistance. *Veterinary parasitology 54*, 1-3 (1994), 259–268.

[83] Rieger, R., Michaelis, A., and Green, M. M. *Glossary of genetics and cytogenetics: classical and molecular*. Springer Science & Business Media, Berlin Germany, 2012.

[84] Roberts, M., and Grenfell, B. The population dynamics of nematode infections of ruminants: periodic perturbations as a model for management. *Mathematical Medicine and Biology: A Journal of the IMA 8*, 2 (1991), 83–93.

[85] Roberts, M., and Grenfell, B. The population dynamics of nematode infections of ruminants: The effect of seasonally in the free-living stages. *Mathematical Medicine and Biology: A Journal of the IMA 9*, 1 (1992), 29–41.

[86] Roberts, M., Smith, G., and Grenfell, B. *Mathematical models for macroparasites of wildlife*, vol. 6. Cambridge University Press, Cambridge UK, 1995.

[87] Rowthorn, R. E., Laxminarayan, R., and Gilligan, C. A. Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface 6*, 41 (2009), 1135–1144.

[88] Sager, S. *Numerical methods for mixed-integer optimal control problems*. Der Andere Verlag, Tönning, 2005.

[89] Segen, J., and Segen, J. *Concise dictionary of modern medicine. New and expanded ed*. McGraw Hill Education, New York, 1992.

[90] Simon, P. L., and Kiss, I. Z. From exact stochastic to mean-field ode models: a new approach to prove convergence results. *The IMA Journal of Applied Mathematics 78*, 5 (2012), 945–964.

[91] Taylor, P., and Lewontin, R. The genotype/phenotype distinction. *Stanford Encyclopedia of Philosophy* (2017).

[92] Travis, J. M., and Dytham, C. The evolution of dispersal in a metapopulation: a spatially explicit, individual-based model. *Proceedings of the Royal Society of London. Series B: Biological Sciences 265*, 1390 (1998), 17–23.

[93] Upton, G., and Cook, I. *A Dictionary of Statistics*. Oxford university press, Oxford UK, 2014.

[94] Viney, M., and Cable, J. Macroparasite life histories. *Current Biology 21*, 18 (2011), R767–R774.

[95] Wächter, A., and Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming 106*, 1 (2006), 25–57.

[96] Waller, P. J. Anthelmintic resistance. *Veterinary parasitology 72*, 3-4 (1997), 391–412.

[97] Wiens, J. A. Metapopulation dynamics and landscape ecology. In *Metapopulation Biology*, I. Hanski and M. E. Gilpin, Eds. Academic Press, San Diego, 1997, pp. 43 – 62.

[98] Wilson, K., Bjørnstad, O., Dobson, A., Merler, S., Poglayen, G., Randolph, S., Read, A., and Skorping, A. Heterogeneities in macroparasite infections: patterns and processes. *The ecology of wildlife diseases 44* (2002), 6–44.

[99] Wilson, K., Grenfell, B. T., and Shaw, D. Analysis of aggregated parasite distributions: a comparison of methods. *Functional Ecology* (1996), 592–601.

[100] Wolstenholme, A. J., Fairweather, I., Prichard, R., von Samson-Himmelstjerna, G., and Sangster, N. C. Drug resistance in veterinary helminths. *Trends in parasitology 20*, 10 (2004), 469–476.

[101] Yusuf, T. T., and Benyah, F. Optimal control of vaccination and treatment for an sir epidemiological model. *World journal of modelling and simulation 8*, 3 (2012), 194–204.

[102] Zhu, C., and Yin, G. On competitive lotka–volterra model in random environments. *Journal of Mathematical Analysis and Applications 357*, 1 (2009), 154–170.

[103] Zur Wiesch, P. A., Kouyos, R., Engelstädter, J., Regoes, R. R., and Bonhoeffer, S. Population biological principles of drug-resistance evolution in infectious diseases. *The Lancet infectious diseases 11*, 3 (2011), 236–247.

Appendices

Appendix A

Chapter 2 Codes

A.1 Isham Codes

A.1.1 Basic Isham Model

The codes to simulate the Isham models from Chapter 2 may be found at:

- Figures 2-1, 2-2:
  - Negative Binomial generating Function $h(z)$: `https://github.com/beth-boulto/Thesis/blob/master/CH2/h_neg_bin.m`
  - Stochastic and Deterministic Simulation: `https://github.com/beth-boulto/Thesis/blob/master/CH2/basic_isham.m`
- Figure 2-3, Treatment of all hosts at a set time: `https://github.com/beth-boulto/Thesis/blob/master/CH2/Isham_time_treat.m`
- Figure 2-4, Treatment of hosts as soon as the host burden exceeds C_max: `https://github.com/beth-boulto/Thesis/blob/master/CH2/Isham_level_timed.m`
- Figure 2-5, Treatment of hosts with burden exceeding C_max at set times: `https://github.com/beth-boulto/Thesis/blob/master/CH2/Isham_time_level_treat.m`

A.2 Larval Model Codes

The code to generate figures 2-7, 2-8, 2-9, 2-10, 2-11, 2-12 and 2-13 may be found at:

- `https://github.com/beth-boulto/Thesis/blob/master/CH2/basic_larval_1.m`

Appendix B

Chapter 3 Codes

B.1 Expected Number of Mated Females

The code to calculate the expected number of females and generate figure 3-2 is located at `https://github.com/beth-boulto/Thesis/blob/master/ch3/single_genotype_per_host.m`

B.2 ODE Models

The codes used to simulate the ODE model in Chapter 3 are given at:

- Function to define (3.8): `https://github.com/beth-boulto/Thesis/blob/master/ch3/mean_field.m`
- Function to define (3.24): `https://github.com/beth-boulto/Thesis/blob/master/ch3/metapop.m`
- Script to use functions to simulate ODE models, calculate equilibrium properties, and plot figures 3-3, 3-5, 3-6 and 3-4: `https://github.com/beth-boulto/Thesis/blob/master/ch3/basic_compare.m`

Appendix C

Chapter 4 Codes

C.1 Bocop Files

The definition files for the continuous optimisation problems for each of the three models can be found at the following locations:

- `https://github.com/beth-boulto/Thesis/tree/master/basicmodel`
- `https://github.com/beth-boulto/Thesis/tree/master/variance_model`
- `https://github.com/beth-boulto/Thesis/tree/master/isham`

These files contain the information that would be input into the BOCOP program when defining a problem and the versions of the files that would be edited to give the correct model dynamics, path conditions, optimisation criteria and dependencies that would need to be written for the problem. Using this information to define a problem using BOCOP should enable the results to be recreated. The instructions to use these files to create the BOCOP executable that may be run are given in the `README.txt` files included in the GitHub folders.

To plot the results of this using Matlab we used the code given at:

- Mean and Variance Models: `https://github.com/beth-boulto/Thesis/blob/master/ch4/bocop_run_no_1.m`
- Isham model: `https://github.com/beth-boulto/Thesis/blob/master/ch4/isham_bocop_process.m`

C.2 Sequentially Determined Integer Controls

This section contains the codes to calculate the integer controls on the three models on sequential sub-intervals. These codes use the functions which define the dynamics of the three models under

treatment, which were given in the previous appendix. The codes which calculate the sequential controls are stored at:

- Figure 4-6, sub-figure (a), (b), (c): https://github.com/beth-boulto/Thesis/blob/master/ch4/sequential/var_inc_integer.m
- Figure 4-6, sub-figure (d) : https://github.com/beth-boulto/Thesis/blob/master/ch4/sequential/isham_integer_final.m

C.3 Brute Force Optimised Integer Controls

The codes in this section are used to calculate the 'optimised 'integer controls for set lengths of sub-intervals for the three models discussed in Chapter 4. The codes may be found at:

- Mean Field Model, figure 4-8 (sub-figure (a)) and 4-9 (sub-figure (a)): https://github.com/beth-boulto/Thesis/blob/master/ch4/brute_f/mean_field_process.m
- Variance Inclusive Model, figure 4-8 (sub-figure (b)) and 4-9 (sub-figure (b)): https://github.com/beth-boulto/Thesis/blob/master/ch4/brute_f/metapop_process.m
- Isham Model, figure 4-8 (sub-figure (c) and (d)) and 4-9 (sub-figure (c)): https://github.com/beth-boulto/Thesis/blob/master/ch4/brute_f/isham_process.m

C.4 Switching System Controls

This section shows some examples of the switching system triggers and simulation codes.

- On switching code for the Mean Field Model when the mean exceeds at set value, C_{max} (figure 4-11): https://github.com/beth-boulto/Thesis/blob/master/ch4/trigger/option_fun_mean_on.m
- Off switching code for when the rate of change in the mean caused by treatment exceeds a set value, $-C$ (figure 4-11) :https://github.com/beth-boulto/Thesis/blob/master/ch4/trigger/option_fun_mean_off.m
- Script which takes on and off switching functions and ODE dynamics function (given previously) and simulates dynamics with switches, figure 4-11, sub-figures (c)and (d)): https://github.com/beth-boulto/Thesis/blob/master/ch4/trigger/mean_f_switcher.m

Similar codes for the variance and Isham models may be found here:

- Variance Model: https://github.com/beth-boulto/Thesis/tree/master/Var_switch
- Isham Model: https://github.com/beth-boulto/Thesis/tree/master/Isham_switch

Appendix D

Chapter 5 Codes

D.1 Haploid Models

This section contains the codes for the Haploid mean field and variance models used in chapter 5. The functions for the model dynamics may be found at:

- Haploid Mean Field Model (6.1): `https://github.com/beth-boulto/Thesis/blob/master/ch5/two_geno/geno_2_mean_dyn_t.m`
- Haploid Variance Inclusive Model (6.2): `https://github.com/beth-boulto/Thesis/blob/master/ch5/two_geno/geno_2_meta_dyn_t.m`

The code which performs basic simulations (e.g figure 5-1) and analytical tests of the equilibrium properties may be found at: `https://github.com/beth-boulto/Thesis/blob/master/ch5/two_geno/basic_analysis.m`. This code makes use of a function to calculate the Jacobian matrix at the non-trivial equilibrium for the variance inclusive model given at `https://github.com/beth-boulto/Thesis/blob/master/ch5/two_geno/jacob_non_triv.m`

D.2 Diploid Model

Similarly to the previous section this section contains the codes for the basic analysis and simulation (figure 5-2) of the Diploid mean field model.

- Dynamics (**??**): `https://github.com/beth-boulto/Thesis/blob/master/ch5/three_geno/geno_3_dyn_t.m`
- Non-Trivial Jacobian Function: `https://github.com/beth-boulto/Thesis/blob/master/ch5/three_geno/Jacob.m`
- Basic analysis and simulation (figure 5-2) script: `https://github.com/beth-boulto/Thesis/blob/master/ch5/three_geno/geno_3_basic.m`

Appendix E

Chapter 6 Codes

E.1 Haploid Model Codes

The code in this section is the Matlab code used to process the results of the BOCOP optimisation of the continuous control on the two genotype models. The BOCOP definition files can be found in the Git repository at the following links.

- Mean field model: `https://github.com/beth-boulto/Thesis/tree/master/two_geno_mean`
- Variance inclusive model: `https://github.com/beth-boulto/Thesis/tree/master/two_geno_var`

To plot the results using Matlab (figure 6-6, 6-7, 6-8) and calculate costs may be found at: `https://github.com/beth-boulto/Thesis/blob/master/ch6/two_geno/bocop_process.m`

E.2 Diploid Model Codes

This section contains the Matlab code to process the continuous optimisation results from the BOCOP codes (definition files at `https://github.com/beth-boulto/Thesis/tree/master/three_geno`) To plot the results using Matlab and calculate costs (figures 6-9 and 6-10) may be found at: `https://github.com/beth-boulto/Thesis/blob/master/ch6/three_geno/bocop_process.m`

Appendix F

Chapter 7 Codes

This appendix contains example codes to sequentially determine an integer control and the optimal integer control on the models shown in chapter 7. These codes may be found at:

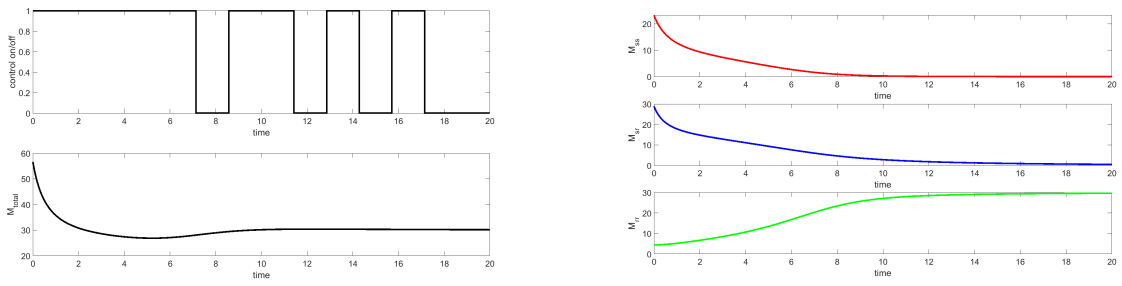
- Code to determine sequential control on Haploid models (figures 7-1, 7-2 and 7-6): `https://github.com/beth-boulto/Thesis/blob/master/ch7/sequential_integer.m`
- Code to calculate the 'optimal 'integer control (figures 7-4, 7-9 and 7-11): `https://github.com/beth-boulto/Thesis/blob/master/ch7/brute_f_integer.m`

These scripts make use of the ODE functions for the models given in a previous appendix. Similar codes were used to calculate the integer controls for the diploid mode but are not given here for the sake of being succinct.

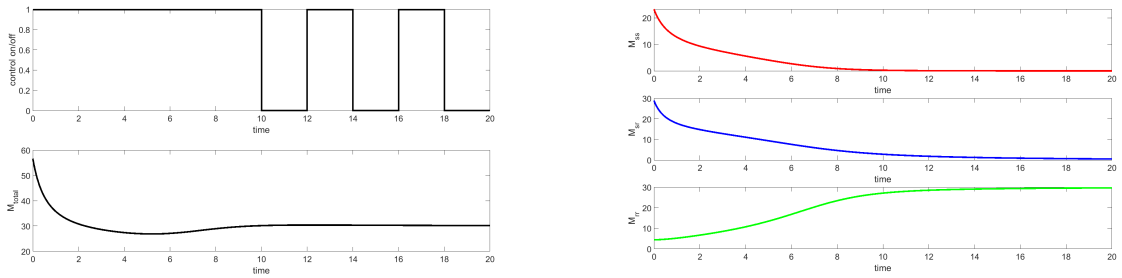
Appendix G

Additional Figures

This appendix contains some additional figures from chapter 7 to allow for comparison with figures contained within the main body of the thesis.

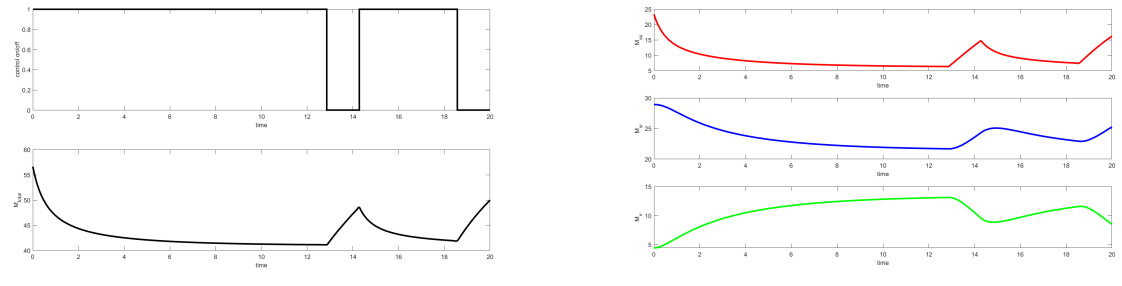


(a) (b)

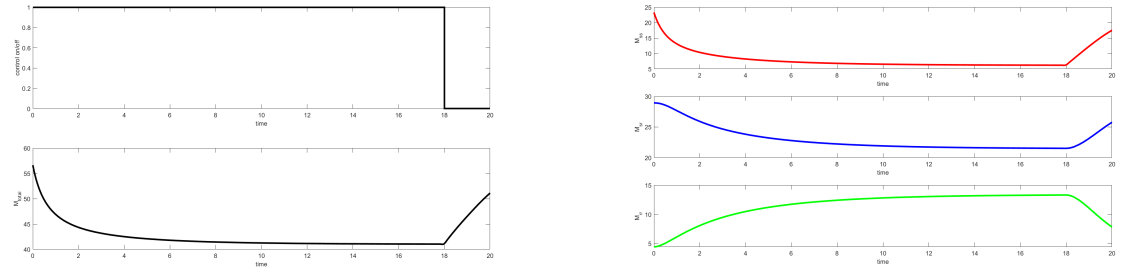


(c) (d)

Figure G-1: Integer control on sub-intervals of length (a)/(b) $\frac{20}{14}$ (c)/(d) 2 optimised across the full interval for the diploid mean field model with heterozygous parasites acting as resistant, $\gamma = 1$. Cost function parameters used for optimisation given by set 1 in table 6.5.



(a) (b)



(c) (d)

Figure G-2: Integer control on sub-intervals of length (a)/(b) $\frac{20}{14}$ (c)/(d) 2 optimised across the full interval for the diploid mean field model with heterozygous parasites acting as resistant, $\gamma = 0$. Cost function parameters used for optimisation given by set 3 in table 6.6.

Appendix H

Future Work Codes

This section contains the code used to model a treatment on the 10% most infected hosts (figure 8-1) in the Isham model for the future work this code may be found at: `https://github.com/beth-boulto/Thesis/blob/master/ch8/isham_timed_level.m`